

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2004, 21:20:15 ; Search time 6482 Seconds

(without alignments)
10939.402 Million cell updates/sec

Title: US-10-018-604-1

Perfect score: 1636

Sequence: 1 aactttttcaatagacaag.....atatgaaaaaaaaaaaaa 1636

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pac:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_rtd:*

37: em_htg_rtd:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1636	100.0	1636	6 A24194	A24194 L.esculentu
2	1636	100.0	1636	6 AR364905	AR364905 Sequence
3	1621	99.1	1621	6 A15981	A15981 L.esculentu
4	1621	99.1	1621	6 AX062336	AX062336 Sequence
5	1621	99.1	1621	6 AX062336	AX062336 Sequence
6	1614	98.6	1614	6 LEPGR	X05656 Tomato mRNA
7	1613.2	98.6	1624	6 I01809	X05656 Tomato mRNA
8	599	34.8	602	8 CPAS05750	AJ505947 Carica pa
9	570	33.6	602	8 LES05947	AJ505947 Lycopersi
10	549.2	33.6	1541	8 AF152758	AF152758 Actinidia
11	525.2	32.1	1595	8 AY043233	AY043233 Vitis vin
12	462	28.2	7456	8 TOMPGAAA	M37304 Tomato poly
13	462	28.2	7456	8 LEPOLYGA	X14074 Tomato gene
14	449.2	27.5	1600	8 AK117942	AK117942 Arabidops
15	445.2	27.2	1296	6 AX432563	AX432563 Sequence
16	445.2	27.2	1296	6 AX651724	AX651724 Sequence
17	445.2	27.2	1296	6 BT005376	BT005376 Arabidops
18	439.8	26.9	1725	8 AVOPOLYCAL	L06094 Avocado pol
19	433	26.5	1755	8 PAPOLYGV	X66426 P.americana
20	431.4	26.4	1364	8 AY078936	AY078936 Arabidops
21	429.6	26.3	1765	8 BNA250919	AJ250919 Brassica
22	428.2	26.2	1767	8 AF062467	AF062467 Cucumis m
23	420.4	25.7	1657	6 A57101	A57101 Sequence 5
24	420.4	25.7	1657	6 AX016328	AX016328 Sequence
25	420.4	25.7	1657	6 BNSAC66	Z49971 B.napus of
26	420.2	25.7	1599	8 BNPGLACR	X95800 B.napus mRN
27	420.2	25.7	1631	6 A62402	A62402 Sequence 1
28	414.6	25.3	1302	8 BRA428543	A428543 Brassica
29	405.8	24.3	1208	8 MAUP14A	L27743 Malus domes
30	397.6	24.3	1208	8 RIATJ147	AJ224147 Rubus ida
31	396.2	24.2	1795	8 AB084461	AB084461 Pyrus com
32	395	24.1	414	8 TOMPGA	M20269 L.esculentu
33	391.6	23.9	1280	6 AX651722	AX651722 Sequence
34	375	22.9	1269	6 AX653992	AX653992 Sequence
35	369	22.6	1182	6 AX653993	AX653993 Sequence
36	331.8	20.3	2207	6 I09042	I09042 Sequence 5
37	331.8	20.3	2207	8 LEPG5	X07410 Lycopersico
38	308.4	18.9	1603	8 AF138858	AF138858 Lycopersi
39	296	18.1	1344	6 AX654086	AX654086 Sequence
40	278.8	17.0	7471	8 FCCCHRAF	L47242 Capsicum an
41	275.8	16.9	573	8 AF152752	AF152752 Actinidia
42	275.8	16.9	667	8 AF152753	AF152753 Actinidia
43	275.2	16.8	690	8 AF152755	AF152755 Actinidia
44	275.2	16.8	696	8 AF152754	AF152754 Actinidia
45	272	16.6	1366	8 AY142668	AY142668 Arabidops

ALIGNMENTS

RESULT 1

A24194 1636 bp DNA linear PAT 04-APR-1995

LOCUS A24194

DEFINITION L.esculentum polygalacturonase clone pTOM6.

ACCESSION A24194

VERSION A24194.1 GI:904404

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

Lycopersicon esculentum (tomato)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;

Asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 1636)

Bridges, L.G., Grierson, D. and Schuch, W.W.

Anti-sense regulation of plant gene expression

ACCESSION AR364905
VERSION AR364905.1 GI:34428082
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1636)
AUTHORS Bridges, I., Schuch, W. and Grierson, D.
TITLE Recombinant DNA containing pectin esterase gene segments
JOURNAL Patent: US 5447867-A 2 05-SEP-1995;
FEATURES
Source Location/Qualifiers
1..1636
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1636; DB 6; Length 1636;
Best Local Similarity 100.0%; Pred. No. 2.9e-246;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATCTTTTCATATGATTAAGTTTAAACCATACATATACATATATCATGTTATCC 60
DB 1 AATCTTTTCATATGATTAAGTTTAAACCATACATATATATCATGTTATCC 60
QY 61 AAAGGAATAGTATCTCTCTCATATATTTTGGCTTCATCATATTTCAACTTGTGAA 120
DB 61 AAAGGAATAGTATCTCTCTCATATATTTTGGCTTCATCATATTTCAACTTGTGAA 120
QY 121 GCATGTTATGATGACATTTATTCAAACAGTTTATGATTAATTTCTTGAACAGAT 180
DB 121 GCATGTTATGATGACATTTATTCAAACAGTTTATGATTAATTTCTTGAACAGAT 180
QY 181 TTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGAGCAAAATATTGAAGCAATA 240
DB 181 TTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGAGCAAAATATTGAAGCAATA 240
QY 241 ATATTTGACAGGTTGATTAATAAAGGATTAAGTATTAATGATTAATGATTAATG 300
DB 241 ATATTTGACAGGTTGATTAATAAAGGATTAAGTATTAATGATTAATGATTAATG 300
QY 301 AGGCTGATGAAAAACATATATATATGATTTGAGCAAGCATGGAATGACATGTT 360
DB 301 AGGCTGATGAAAAACATATATATATGATTTGAGCAAGCATGGAATGACATGTT 360
QY 361 CATCTAGAACCTGTTCAATTTTGCTTAAACCAAGATTAATCTTCAAGCAAA 420
DB 361 CATCTAGAACCTGTTCAATTTTGCTTAAACCAAGATTAATCTTCAAGCAAA 420
QY 421 TCACCTTTTCAAGTTCATGAGATCTTCTATTTCAATTAAGATTTTGGATCTTAAAG 480
DB 421 TCACCTTTTCAAGTTCATGAGATCTTCTATTTCAATTAAGATTTTGGATCTTAAAG 480
QY 481 CATCTAGAAATTTTCAAGTTCATGAGATCTTCTATTTCAATTAAGATTTTGGATCTT 540
DB 481 CATCTAGAAATTTTCAAGTTCATGAGATCTTCTATTTCAATTAAGATTTTGGATCTT 540
QY 541 AAAATTTAGTGTGAGAGAGAGAACTATCAATGCAATGCAATGATGATGAGCAAA 600
DB 541 AAAATTTAGTGTGAGAGAGAGAACTATCAATGCAATGCAATGATGATGAGCAAA 600
QY 601 GTTCTTGCAAAATTAATCACTGCAATGCAAGGATGCAACAGGCTTAACTTTCT 660
DB 601 GTTCTTGCAAAATTAATCACTGCAATGCAAGGATGCAACAGGCTTAACTTTCT 660
QY 661 GGAATTTGAAAAATTTGAAGTAACTTAAGATTAAGTAAAGTAAAGTAAAGTAAAGT 720
DB 661 GGAATTTGAAAAATTTGAAGTAACTTAAGATTAAGTAAAGTAAAGTAAAGTAAAGT 720
QY 721 TCAATTTGAGTCACTATGTTGAGCTTCAATTTGATGATGATGATGATGATGATG 780
DB 721 TCAATTTGAGTCACTATGTTGAGCTTCAATTTGATGATGATGATGATGATGATG 780
QY 781 AGAGCCCAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 840

DB 781 AGAGCCCAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 CTATTTATGGAACAGTGTATGATTTCAATTTGTTCTGATCTCAAAATGTCAGG 900
DB 841 CTATTTATGGAACAGTGTATGATTTTCAATTTGTTCTGATCTCAAAATGTCAGG 900
QY 901 CCACAAATATTAATCTTGTGTCAGGATGATGATTAATGATTAAGATCTAGATCTGGA 960
DB 901 CCACAAATATTAATCTTGTGTCAGGATGATGATTAATGATTAAGATCTAGATCTGGA 960
QY 961 ATTGAGAGCTTATGTTCTTAATGTTTCTGTAATTAAGTAAATGAAAGTAAATTAATG 1020
DB 961 ATTGAGAGCTTATGTTCTTAATGTTTCTGTAATTAAGTAAATGAAAGTAAATTAATG 1020
QY 1021 ATGAGTTAGATCAAGATCTTGAGAGAGATCTGAGCAAGTCAAGTCAAGTCAAGTCT 1080
DB 1021 ATGAGTTAGATCAAGATCTTGAGAGAGATCTGAGCAAGTCAAGTCAAGTCAAGTCT 1080
QY 1081 TGAATGTGAAATGCAAGAGCTTAATGATCCATTAATTAATGACCAAAATCTAATGATC 1140
DB 1081 TGAATGTGAAATGCAAGAGCTTAATGATCCATTAATTAATGACCAAAATCTAATGATC 1140
QY 1141 GAGTTGAACCATGATTAACACAGTTTTCAGAGTTCAAGTAAATTAATGATTAATG 1200
DB 1141 GAGTTGAACCATGATTAACACAGTTTTCAGAGTTCAAGTAAATTAATGATTAATG 1200
QY 1201 ATATCAAGGAGCAAGTCAAGCAAGGATGAGCAATTAATGATGAGCAAGCAAGTCT 1260
DB 1201 ATATCAAGGAGCAAGTCAAGCAAGGATGAGCAATTAATGATGAGCAAGCAAGTCT 1260
QY 1261 CATGTGAAGGATTAATTAATGAGATTAATTAATTAATGAGGAGAAAGTGAAGAAC 1320
DB 1261 CATGTGAAGGATTAATTAATGAGATTAATTAATTAATGAGGAGAAAGTGAAGAAC 1320
QY 1321 AGGCTAGTGAAGAAATGTCATTTTAACATGCTGAACATGTTACACCAAGTCACTT 1380
DB 1321 AGGCTAGTGAAGAAATGTCATTTTAACATGCTGAACATGTTACACCAAGTCACTT 1380
QY 1381 CACTAGAAATTTGAGAGATGAGCTCTTTGATTAATTAATTAATTAATTAATTAAT 1440
DB 1381 CACTAGAAATTTGAGAGATGAGCTCTTTGATTAATTAATTAATTAATTAATTAAT 1440
QY 1441 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
DB 1441 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
QY 1501 TTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
DB 1501 TTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
QY 1561 GTCAAAAGTTGAGATGATCTTTTATGATTAATTAATTAATTAATTAATTAAT 1620
DB 1561 GTCAAAAGTTGAGATGATCTTTTATGATTAATTAATTAATTAATTAATTAAT 1620
QY 1621 GAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1636
DB 1621 GAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1636
RESULT 3
A15981
LOCUS
DEFINITION
ABSTRACT
VERSION
KEYWORDS
ORGANISM
REFERENCE
1 (bases 1 to 1621)
1621 bp RNA linear PAT 02-MAR-1994
Lycopersicon esculentum
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

RESULT 4
AX062336

Applicant

ORIGIN

Query Match 1621	99.1%	Score 1621	DB 6	length 1621
Best Local Similarity	100.0%	Pred. Nc	6.4e-244	
Matches 1621	0	Mismatches	0	Gaps 0

QY	1	AATCTTTTCAATAGACAAGTTTAAAAACCAACATATATTAACAATATACAGTTATCC	60
Db	1	AATCTTTTCAATAGACAAGTTTAAAAACCAATCAATATTAACAATATACAGTTATCC	60
QY	61	AAAGGAATAGATATTCCTCTCATTTATTTTGGCTTCATCATATTTCAACTTGTAGAA	120
Db	61	AAAGGAATAGATATTCCTCTCATTTATTTTGGCTTCATCATATTTCAACTTGTAGAA	120
QY	121	GCAATGTTATTCATGACAAATTATTTCAACAAGTTATGATAAATTTCTTGAACAAGAT	180
Db	121	GCAATGTTATTCATGACAAATTATTTCAACAAGTTATGATAAATTTCTTGAACAAGAT	180
QY	181	TTGCTCATGATTTTCAAGCTTATCTTCTATTTTAGCAAAAATAATTGAACCAACATA	240
Db	181	TTGCTCATGATTTTCAAGCTTATCTTCTATTTTAGCAAAAATAATTGAACCAACATA	240
QY	241	ATATTGACAAGTTGATATAAAAAATGGATTTAAAGTGATTATGACTTACCTTTGAGCTA	300
Db	241	ATATTGACAAGTTGATATAAAAAATGGATTTAAAGTGATTATGACTTACCTTTGAGCTA	300
QY	301	AGGCGATGAGAAAAACATATGATATATTTGCAATTTGAGCAACATGAGATGAAGATGT	360
Db	301	AGGCGATGAGAAAAACATATGATATATTTGCAATTTGAGCAACATGAGATGAAGATGT	360
QY	361	CATCTAGAAACACCTGTTCAATTGTGGTTCCTATAAAAAACAAGATTAATCTTCAAGCAAA	420
Db	361	CATCTAGAAACACCTGTTCAATTGTGGTTCCTATAAAAAACAAGATTAATCTTCAAGCAAA	420
QY	421	TCACCTTTTCAGTCCATGACAGATCTTCTATTTCAGTAAAGAATTTTGGATCCTTGAAG	480

[illegible]

QY 1561 GTCAAAAGTTGACGATTTACTTTTAAATGACAAAATAATPAAATGGTATTATAT 1620
| | | | |
Db 1561 GTCAAAAGTTGACGATTTACTTTTAAATGACAAAATAATPAAATGGTATTATAT 1620
| | | | |
QY 1621 G 1621
|
Db 1621 G 1621
|
RESULT 5
LEPG2AR 1621 bp mRNA linear PLN 27-MAR-1995
LOCUS
DEFINITION Tomato mRNA for polygalacturonase-2a (PG-2a).
ACCESSION X04583
VERSION X04583.1 GI:19291
KEYWORDS glycoprotein; polygalacturonase; polygalacturonase-2a; signal peptide;
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1621)
Grierson, D., Tucker, G.A., Keen, J., Ray, J., Bird, C.R. and Schuch, W. Sequencing and identification of a cDNA clone for tomato polygalacturonase
Nucleic Acids Res. 14 (21), 8595-8603 (1986)
87066731
3786135
The predicted N-terminal AA sequence of PG-2a occurs at AA72. Since the ATG at bp 51 codes for the only inframe methionine preceding the N-terminal sequence of PG-2a, it is probable that the mRNA is translated with a 71AA pre-sequence that is subsequently cleaved to give a 41,828D mature protein.
Data kindly reviewed (22-APR-1987) by D. Grierson.
Location/Qualifiers
1..1621
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/strain="Alisa Craig"
/db_xref="taxon:4081"
/clone="PTOM 6"
/tissue_type="fruit"
/dev_stage="fruit ripe"
51..1424
/note="unnamed protein product; precursor polypeptide (AA -71 to 386)"
/codon_start=1
/protein_id="CAA28254.1"
/db_xref="GI:19292"
/db_xref="GOA:P05117"
/db_xref="SWISS-PROT:P05117"
/translation="WVIGRNSILILIIIFASISITCRSNVIDNLFKQVYDNIIIEQF AHDFOAYLSYLSKINISNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAEQAMNEA CSSRTPQVFPVKINKNYLLKQITFSGPSSISVKIPSLASSKISDYKRLMIAR DSVQNLVVGSGGTINGNGOVMPSSCKINKSLPCDAPFTALTPWCKLTKNNLSKX AAOIHKESSCTNVVANSMLTNASAKSPRTGCVHNSNVOYQIISTITIGTDDCSISV SGGSNVQATNTITCGPBGHISTIGLSGNSSEAVSNVTNEAKITGAENGVIKTMQGG SGQASNIFLNVEMQDVAFPIIIDQNYCDREVPICIQPSAVQKNVVENKIGTSATK VAIKFDCTNPFCEGIIIMENINLVGSEKPESEATCKNVHFNNAEHVTCHSLEISD EALILNY"
264..1421
/product="mature polygalacturonase (AA 1-386)"
51..263
/note="put. pre-peptide (AA -71 to -1)"
615..623
/note="pot. N-glycosylation site"
768..776
/note="pot. N-glycosylation site"
906..914
/note="pot. N-glycosylation site"
misc_feature
981..989

QY 1 AATCTTTCAATAGACAAAGTTTAAAAACCATACATATACATATATCATGTTATCC 60
| | | | |
Db 1 AATCTTTCAATAGACAAAGTTTAAAAACCATACATATACATATATCATGTTATCC 60
| | | | |
QY 61 AAAGGATATGATTTCTCTTCATATTTATTTTCTTCATCATATTTCACTTTAGAA 120
| | | | |
Db 61 AAAGGATATGATTTCTCTTCATATTTATTTTCTTCATCATATTTCACTTTAGAA 120
| | | | |
QY 121 GCAATGTTATGATGACATTTATTCAAACAGTTATATATATATATATATATATAT 180
| | | | |
Db 121 GCAATGTTATGATGACATTTATTCAAACAGTTATATATATATATATATATATAT 180
| | | | |
QY 181 TTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGACAAAATATTTGAAGCAACATA 240
| | | | |
Db 181 TTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGACAAAATATTTGAAGCAACATA 240
| | | | |
QY 241 ATATTGACAAAGTTGATATAAATGGATTAAGTGATATATGTAAGCTTTGAGCTTA 300
| | | | |
Db 241 ATATTGACAAAGTTGATATAAATGGATTAAGTGATATATGTAAGCTTTGAGCTTA 300
| | | | |
QY 301 AGGGTGATGAAAAACATATGATATATGATTCATTTGAGCAAGCATGAAATGAAAGATGT 360
| | | | |
Db 301 AGGGTGATGAAAAACATATGATATATGATTCATTTGAGCAAGCATGAAATGAAAGATGT 360
| | | | |
QY 361 CATCTAGAACACCTGTTCATTTTGTGTTCCCTTAAACAAAGATTTATCTTCAAGCAAA 420
| | | | |
Db 361 CATCTAGAACACCTGTTCATTTTGTGTTCCCTTAAACAAAGATTTATCTTCAAGCAAA 420
| | | | |
QY 421 TCACCTTTTCAGGTCATGACAGATCTTCTATTTCAAGTAAAGATTTTGTGATCTTGAAG 480
| | | | |
Db 421 TCACCTTTTCAGGTCATGACAGATCTTCTATTTCAAGTAAAGATTTTGTGATCTTGAAG 480
| | | | |
QY 481 CATCTAGTAAATTTTCAGCTCAACAAGATGAAAGGCTTTGATGCTTTGATAGGCTC 540
| | | | |
Db 481 CATCTAGTAAATTTTCAGCTCAACAAGATGAAAGGCTTTGATGCTTTGATAGGCTC 540
| | | | |
QY 541 AAAATTTAGTTGTGAGAGAGAGAACTATCATATGCAATGCAATGCAAGTATGAGGCCAA 600
| | | | |
Db 541 AAAATTTAGTTGTGAGAGAGAGAACTATCATATGCAATGCAATGCAAGTATGAGGCCAA 600
| | | | |
QY 601 GTTCTTGCAAAATTAATATATCTGCCATGACGAGATGACCAACGCGCTTAACCTTCT 660
| | | | |
Db 601 GTTCTTGCAAAATTAATATATCTGCCATGACGAGATGACCAACGCGCTTAACCTTCT 660
| | | | |
QY 661 GGAATTTGCAAAATTTGAAAGGAAATATCTAAAGATGAAATGCAACAAATTTCTATA 720
| | | | |
Db 661 GGAATTTGCAAAATTTGAAAGGAAATATCTAAAGATGAAATGCAACAAATTTCTATA 720
| | | | |
QY 721 TCAATTTGAGTCATGACATATATGTTAGCTTCAATTTTGTGATCATGCTTCAGCAA 780
| | | | |
Db 721 TCAATTTGAGTCATGACATATATGTTAGCTTCAATTTTGTGATCATGCTTCAGCAA 780
| | | | |
QY 781 AGAGGCCAAATATCATGATGAGTCCATGATCAAAATGCAATATATTTCAAAATATCTGATA 840
| | | | |
Db 781 AGAGGCCAAATATCATGATGAGTCCATGATCAAAATGCAATATATTTCAAAATATCTGATA 840
| | | | |
QY 841 CTATATTTGGAACAGGATATGATTTTCAATTTGTTTGGATCTCAAAATGAGCAGG 900
| | | | |
Db 841 CTATATTTGGAACAGGATATGATTTTCAATTTGTTTGGATCTCAAAATGAGCAGG 900
| | | | |
QY 901 CCACAAATATTTACTTGTGTCAGGTCATGATATAGTATGGAAGCTTAGGATCTGAAA 960
| | | | |

D	b		484	TAATAATTTCAAGCTTCACAAAGATAGAAAGCTTTGAAATGCTTTTGATAGTGTTCAAAATTT	543
Q	y		548	AGTTGTGGAGAGAGGAACCTATCAATGGCAATGACAAAGTATGAGGCAAGTCTTG	607
D	b		544	AGTTGTTGAGAGAGGAGAACTATCAATGGCAATGACAAAGTATGAGTGTGGCAAGTCTTG	603
Q	y		608	CAAAATAATAATATCATCTGCCATGAGAGGATGACCAACGCGCTTAACTTCTGGAATTG	667
D	b		604	CAAAATAATAATAATCACTGCCATGAGAGGATGACCAACGCGCTTAACTTCTGGAATTG	663
Q	y		668	CAAAATAATGAAAGTAAATCTTAAAGTAAAAATGACAAACAAATTCATATGAAAT	727
D	b		664	CAAAATAATGAAAGTAAATCTTAAAGTAAAAATGACAAACAAATTCATATGAAAT	723
Q	y		728	TGAATCATGCACTAATGTTAGCTTCAAAATTTGATGATCAATGCTTGAGCAAAAGGCC	787
D	b		724	TGAGTCATGCACTAATGTTAGCTTCAAAATTTGATGATCAATGCTTGAGCAAAAGGCC	783
Q	y		788	AAATAATCTATGAGTCCATGTTATCAAAATCTCAATATATTCAAATATCTGATCTATAT	847
D	b		784	AAATAATCTATGAGTCCATGTTATCAAAATCTCAATATATTCAAATATCTGATCTATAT	843
Q	y		848	TGGAACAGGTATGATGTTGATTTCAATGTTCTGAGACTCAAAATGAGGCCACAA	907
D	b		844	TGGAACAGGTATGATGTTGATTTCAATGTTCTGAGACTCAAAATGAGGCCACAA	903
Q	y		908	TATTAATCTTGAGTCCAGGTCAATGTTATAGTATTTGGAAGCTTAGATCTGAAATTCAGA	967
D	b		904	TATTAATCTTGAGTCCAGGTCAATGTTATAGTATTTGGAAGCTTAGATCTGGAATTCAGA	963
Q	y		968	AGCTTAATGTGTCTAATGTTACTGTATATGAAGCCAAATATCGGTGCCAAATGAGAGT	1022
D	b		964	AGCTTAATGTGTCTAATGTTACTGTATATGAAGCCAAATATCGGTGCCAAATGAGAGT	1022
Q	y		1028	TAGATATCAAGCTTGCGAGGAGATCTGGAACAAGCTGCAACATCAATTTCTGAATGT	1088
D	b		1024	TAGATATCAAGCTTGCGAGGAGATCTGGAACAAGCTGCAACATCAATTTCTGAATGT	1083
Q	y		1088	GGAATGCAACAGCTTAAGTATCCCATATATTAAGACCAAAACATTTGTGATCGAGTTGA	1147
D	b		1084	GGAATGCAACAGCTTAAGTATCCCATATATTAAGACCAAAACATTTGTGATCGAGTTGA	1147
Q	y		1148	ACCATGTTATCAACAGTTTTCAAGCATTTCAAGTGAATAATGAGTGTGATGAGATATCAA	1207
D	b		1144	ACCATGTTATCAACAGTTTTCAAGCATTTCAAGTGAATAATGAGTGTGATGAGATATCAA	1203
Q	y		1208	GGGCAACAGTGCACAAAGGTGGCCATATAATTTGATTTGCAAGCAAACTTTCCATGTGA	1267
D	b		1204	GGGCAACAGTGCACAAAGGTGGCCATATAATTTGATTTGCAAGCAAACTTTCCATGTGA	1263
Q	y		1268	AGGAATTTATATGAGAAATATAAATTTAGTAGGGGAAAGTGGAAAACATCAGAGGCTAC	1322
D	b		1264	AGGAATTTATATGAGAAATATAAATTTAGTAGGGGAAAGTGGAAAACATCAGAGGCTAC	1322
Q	y		1328	GTGCAAAAATGTCCATTTTAAACAATGCTGAACATGTTACACCACTGCACTTCACTAGA	1387
D	b		1324	GTGCAAAAATGTCCATTTTAAACAATGCTGAACATGTTACACCACTGCACTTCACTAGA	1383
Q	y		1388	AATTTGAGAGATGAAGCTCTTTTGATATAATTAATTAATTTATATATCTATGATCTTGAAT	1444
D	b		1384	AATTTGAGAGATGAAGCTCTTTTGATATAATTAATTAATTTATATATCTATGATCTTGAAT	1444
Q	y		1448	ATAGCAGATATGATATATACAAATTAACAAATCTATCTATGATGATTAATTAATTAAT	1507
D	b		1444	ATAGCAGATATGATATATACAAATTAACAAATCTATCTATGATGATTAATTAATTAAT	1507
Q	y		1508	TAAATATGATCGAATGGAATTTTATAATAGACTATGATTAATTTCTAATTTCTAGCAAAA	1567
D	b		1504	TAAATATGATCGAATGGAATTTTATAATAGACTATGATTAATTTCTAATTTCTAGCAAAA	1567
Q	y		1568	GTTTGACGATTTGATCTTTTATATGACAAAAATAAATGTTATTTATATG 1621	

DB	1564	GTTCGACATGTGCTACTTTTATATGTACAAAATAATATAAGTTATTATATG	1617
RESULT 7			
LOCUS	101809	1624 bp ss-DNA	linear
DEFINITION	Sequence 1 from Patent US 4801540.		PAT 21-MAY-1992
ACCESSION	101809		
VERSION	101809.1	GI:269802	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1624)		
TITLE	Hiatc,W.R., Sheehy,R.E., Shewmaker,C.K., Kridl,J.C. and Knauft,V.		
JOURNAL	PG gene and its use in plants		
FEATURES	Patent: US 4801540-A 1 31-JAN-1989;		
source	Calgene, Inc.; Davis, CA		
	Location/Qualifiers		
	1..1624		
ORIGIN	/organism="unknown"		
	/mol_type="unassigned DNA"		
Query Match	98.6%;	Score 1613.2;	DB 6;
Best Local Similarity	99.8%;	Fred. No. 1.1e-242;	Length 1624;
Matches 1615;	Conservative 0;	Mismatches	Indels 0; Gaps 0;
QY	8	TTCAATAAGACAGTTTAAAAACCATACATCATTAACATATATCATGTTATCCAAAGGAA	67
DB	4	TTCAATAGACACAGTTTAAAAACCATACATCATTAACATATATCATGTTATCCAAAGGAA	63
QY	68	TAGTATTCCTCTTCATTAATTTTGGCTTCATCAATTTCAACTGTAGAACATGT	127
DB	64	TAGTATTCCTCTTCATTAATTTTGGCTTCATCAATTTCAACTGTAGAACATGT	123
QY	128	TATGATGACAAATTTATCCAAACAGTTTATGATATATTTCTTGAAACAAGATTTGCTCA	187
DB	124	TATGATGACAAATTTATCCAAACAGTTTATGATATATTTCTTGAAACAAGATTTGCTCA	183
QY	188	TGATTTTCAACCTTATCTTCTTATTTAGACAAAATTTGAAAGCAACATTAATTTGA	247
DB	184	TGATTTTCAACCTTATCTTCTTATTTAGACAAAATTTGAAAGCAACATTAATTTGA	243
QY	248	CAAGGTTGATAAAAATGGATTTAAAGTATTAATGTACTTGAAGCTTAAAGGTGA	307
DB	244	CAAGGTTGATAAAAATGGATTTAAAGTATTAATGTACTTGAAGCTTAAAGGTGA	303
QY	308	TGAAAAACATATGATATAATTTGATTTAGACAAAGCATGAAATGACATGTTCACTTGA	367
DB	304	TGAAAAACATATGATATAATTTGATTTAGACAAAGCATGAAATGACATGTTCACTTGA	363
QY	368	AACACCTGTTCAATTTGGTTCCTTAAACAAAGATATATCTTCAAGCAAAATCACTT	427
DB	364	AACACCTGTTCAATTTGGTTCCTTAAACAAAGATATATCTTCAAGCAAAATCACTT	423
QY	428	TTCAAGTTCATGACATCTTCTATTTTCAGTAAAGATTTTGGATCCTTGAAGCATCTAG	487
DB	424	TTCAAGTTCATGACATCTTCTATTTTCAGTAAAGATTTTGGATCCTTGAAGCATCTAG	483
QY	488	TAAATTTTCAAGCTTCAAAAGATAGAAGCTTTGGATTTGTTGATGATGTTCAAAATTT	547
DB	484	TAAATTTTCAAGCTTCAAAAGATAGAAGCTTTGGATTTGTTGATGATGTTCAAAATTT	543
QY	548	AGTTGTTGAGAGAGAGAACTATCAATGGCAATGGCAAACTATGAGTGGCCAAAGTCTTG	607
DB	544	AGTTGTTGAGAGAGAGAACTATCAATGGCAATGGCAAACTATGAGTGGCCAAAGTCTTG	603
QY	608	CAAAATTAATTAATTCATGTCATGACAGGATGACCAACGACCTTAACTTCTGGAATTTG	667
DB	604	CAAAATTAATTAATTCATGTCATGACAGGATGACCAACGACCTTAACTTCTGGAATTTG	663
QY	668	CAAAATTTTGAAGTGAATTAATCTTAAGAGTTAAATAATGACAAACATTTCAATCAAAAT	727


```

Db      664 CAAAAATTGAAAGTGAATATCTAAGAGTAAAGCAACCAATTATATCAAAATT 723
      728 TGAATCATGCACTAATGTTAGCTTCAATTTGATGATCAATAGCTTCAGCAAGAGCCC 787
      724 TGAGTCATGCACTAATGTTAGCTTCAATTTGATGATCAATAGCTTCAGCAAGAGCCC 783
      788 AAATACTGATGAGTCCATGATCAATTAATCAATTAATCAATTAATCAATTAAT 847
      784 AAATACTGATGAGTCCATGATCAATTAATCAATTAATCAATTAATCAATTAAT 843
      848 TGAACACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
      844 TGAACACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 903
      908 TATTACTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 967
      904 TATTACTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 963
      968 AGCTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
      964 AGCTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023
      1028 TAGATCATGACCTTGGCAGGAGGATGATGATGATGATGATGATGATGATGAT 1087
      1024 TAGATCATGACCTTGGCAGGAGGATGATGATGATGATGATGATGATGATGAT 1083
      1088 GGAATATGCAAGCGTTAGTATCCATTAATTAATTAATTAATTAATTAATTA 1147
      1084 GGAATATGCAAGCGTTAGTATCCATTAATTAATTAATTAATTAATTAATTA 1143
      1148 ACCATGATATCAACAGTTTTCAGCAGTTTCAGCAGTTTCAGCAGTTTCAGC 1207
      1144 ACCATGATATCAACAGTTTTCAGCAGTTTCAGCAGTTTCAGCAGTTTCAGC 1203
      1208 GGGCACAAGTGCACAAAGGTGGCCTAAATTTGATGACACAACTTTCAGTGA 1267
      1204 GGGCACAAGTGCACAAAGGTGGCCTAAATTTGATGACACAACTTTCAGTGA 1263
      1268 AGGAATATATAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1327
      1264 AGGAATATATAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1323
      1328 GTGCACAAAATGTCATTTTAACATGCTGAACATGTTACACACACTGCATCTA 1387
      1324 GTGCACAAAATGTCATTTTAACATGCTGAACATGTTACACACACTGCATCTA 1383
      1388 AATTTCAGAGGATGAGAGCTCTTTTGTATTAATTAATTAATTAATTAATTA 1447
      1384 AATTTCAGAGGATGAGAGCTCTTTTGTATTAATTAATTAATTAATTAATTA 1443
      1448 ATGACAGATGATGATATATCAATTAATTAATTAATTAATTAATTAATTAAT 1507
      1444 ATGACAGATGATGATATATCAATTAATTAATTAATTAATTAATTAATTAAT 1503
      1508 TAAATATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1567
      1504 TAAATATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
      1568 GTTTGACGATTTGATCTTTTAAATGATCAAAAATTAATTAATTAATTAATTA 1625
      1564 GTTTGACGATTTGATCTTTTAAATGATCAAAAATTAATTAATTAATTAATTA 1621

```

RESULT 8
 CPAS05750 599 bp mRNA linear PLN 04-SEP-2003
 LOCUS
 DEFINITION Carica papaya partial mRNA for putative polygalacturonase (pg
 gene).
 ACCESSION AF505750
 VERSION AF505750.1 GI:34481836
 KEYWORDS Pg gene; polygalacturonase.
 SOURCE Carica papaya (papaya)

```

ORGANISM Carica papaya
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Caricaceae; Carica.
AUTHORS Saiprasad, G.V.S. and Lalitha, A.
TITLE Isolation, cloning and characterization of polygalacturonase gene
JOURNAL from ripening papaya fruit (Carica papaya cv. Singapore pp-1)
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 599)
JOURNAL Saiprasad, G.V.S.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2002) Saiprasad G.V.S., Biotechnology, Indian
Institute of Horticultural Research, IIHR, Hesaraghatta lake post,
Bangalore, Karnataka, INDIA
FEATURES
source
1..599
/organism="Carica papaya"
/mol_type="mRNA"
/cultivar="Singapore pp-1"
/db_xref="taxon:3649"
/tissue_type="fruit"
1..599
/gene="pg"
<1..>599
/gene="pg"
/codon_start=2
/product="putative polygalacturonase"
/protein_id="CAD44275.1"
/db_xref="GI:34481837"
/translation="EQAMNDCSRIPVQFVPERKNNYLIKQITPSPGSSISVKIF
GSLKASKIDYDRLMLAFDSVOMLVGGGTIGNGGVWVPPSSCKTKNSLPCRD
PTALPTWCKNLKVNLSKNAQOIHKEPSCNNVVASNIMINIASAKSPMTDGVHSN
TYIQISDRIIGNDICISIVSGSNVQATNIRICPGHG"
ORIGIN
Query Match 36.6%; Score 599; DB 8; Length 599;
Best Local Similarity 100.0%; Pred. No. 3,9e-84;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db      335 TGAACAGCATGGAAGAAGCAATGTTCACTTAAGAACCGTTCATTTGGTCTCTAA 394
      1 TGAACAGCATGGAAGAAGCAATGTTCACTTAAGAACCGTTCATTTGGTCTCTAA 60
      395 AAACAGAAATTAATCTTCTCAGCAAAATCACCTTTTCAGTGCATGAGATCTTCTATTC 454
      61 AAACAGAAATTAATCTTCTCAGCAAAATCACCTTTTCAGTGCATGAGATCTTCTATTC 120
      455 AGTAAAGATTTTGGATCCTTAAGAGATCTAGTAAATTTTCAGATCAAGATGAG 514
      121 AGTAAAGATTTTGGATCCTTAAGAGATCTAGTAAATTTTCAGATCAAGATGAG 180
      515 GCTTGGATGCTTTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
      181 GCTTGGATGCTTTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
      575 TGGCAATGACAGATAGTGTGCGCAAGTTCTTGCAAAATTAATTAATTAATTAATTA 634
      241 TGGCAATGACAGATAGTGTGCGCAAGTTCTTGCAAAATTAATTAATTAATTAATTA 300
      635 GGATGACCAACGAGCCTTAACCTTTCTGGAATTTGCAAAAATTTGAAGTGAATATCT 694
      301 GGATGACCAACGAGCCTTAACCTTTCTGGAATTTGCAAAAATTTGAAGTGAATATCT 360
      695 GAGTAAATAATGCAACAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 754
      361 GAGTAAATAATGCAACAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
      755 AAATTTGATGATCAATGCTTCAGCAAAAGAGCCCAATTAATTAATTAATTAATTA 814
      421 AAATTTGATGATCAATGCTTCAGCAAAAGAGCCCAATTAATTAATTAATTAATTA 480
      815 TACTCAATATTAATCAAAATATCTGATCTAATTAATTTGCAACGAGTATGATTTCAAT 874

```


Db 481 TACTCAATATATCAATATATCTGATATCTATTTATGAGCAGGTATGATGATTTCAAT 540
QY 875 TGTTCCTGATCTCAAAATGTCGAGCCCAAAATATTAATTGTCGATCATGTA 933
Db 541 TGTTCCTGATCTCAAAATGTCGAGCCCAAAATATTAATTGTCGATCATGTA 599
RESULT 9
LOCUS LES505947 602 bp mRNA linear PLN 04-SEP-2003
DEFINITION Lycopersicon esculentum partial mRNA for putative polygalacturonase
(pg gene) .
ACCESSION AJ505947
VERSION AJ505947.1 GI:34481842
KEYWORDS pg gene; polygalacturonase.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 Saiprasad,G.V.S.
AUTHORS Isolation, Cloning and characterization of polygalacturonase gene
TITLE from fruit tissue of Lycopersicon esculentum cv. Arka vikas
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 602)
AUTHORS Saiprasad,G.V.S.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-2002) Saiprasad G.V.S., Biotechnology, Indian
Institute of Horticultural Research, IHR, Hesaraghatta Lake post,
Bangalore, Karnataka, INDIA
FEATURES
source Location/Qualifiers
1..602
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultiivar="Arka vikas"
/db_xref="taxon:4081"
/tissue_type="fruit"
1..602
/gene="pg"
/gene="pg"
/gene="pg"
/codon_start=2
/evidence="experimental"
/product="putative polygalacturonase"
/protein_id="CAD44521.1"
/db_xref="GI:34481843"
/translation="EQAMNEACSSRTPVQFVVPKKNKYLKQITFSGFGRSSISEKIF
GSLFASKISDYKDRILDDCLKSVQNLVVGSGGTLNNGGVMPSSCKIKRSLPCRD
APTALTFFNCKLKVNNLKSNAQOIHKFESCTNVVNSNIMINASAKSPNTDGVHS
NTQYIQISDTITIGTDCISIVSGSNVQAINITGPHG"
ORIGIN
Query Match 34.8%; Score 570; DB 8; Length 602;
Best Local Similarity 97.8%; Pred. No. 1.3e-79;
Matches 589; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
QY 335 TGAGCAAGCATGGAATGAAGATGTTCACTAGAAACACCTGTTCAATTTGGTCTCTAA 394
Db 1 TGAGCAAGCATGGAATGAAGATGTTCACTAGAAACACCTGTTCAATTTGGTCTCTAA 60
QY 395 AAACAGAAATATCTCTCAAGCAAAATCACCTTTCAAGTCATGCAATCTTCATATTC 454
Db 61 AAACAGAAATATCTCTCAAGCAAAATCACCTTTCAAGTCATGCAATCTTCATATTC 120
QY 455 AGTAAAGATTTTGGATCCTTAGAAGCATCTAGTAAATTTCAAGCTACAAAGATAGAAG 514
Db 121 AGAAAGATTTTGGATCCTTAGAAGCATCTAGTAAATTTCAAGCTACAAAGATAGAAG 180
QY 515 GCTTTGGATT--GCTTTGATAGTGTTCAAAATTTAGTGTGGAGAGAGAACTAT 571
Db 181 GCTTTGGATGCTTTGAAAAGTGTTCAAAATTTAGTGTGGAGAGAGAACTAT 240

QY 572 CAATGCAATGACAAAGATATGTTGGCCAAAGTTCTTGCAAAAATAAATATCACTGCCATG 631
Db 241 CAATGCAATGACAAAGATATGTTGGCCAAAGTTCTTGCAAAAATAAATATCACTGCCATG 300
QY 632 CAGGATGACCAACAGGCTTACCTTCTGAAATGCAAAAATTTGAAAGTAATATCT 691
Db 301 CAGGATGACCAACAGGCTTACCTTCTGAAATGCAAAAATTTGAAAGTAATATCT 360
QY 692 AAAGCTTAAATGACAAACAAATTCATATCAAAATTTGATGATGATGATGATGATG 751
Db 361 AAAGCTTAAATGACAAACAAATTCATATCAAAATTTGATGATGATGATGATGATG 420
QY 752 TTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 811
Db 421 TTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
QY 812 AATATCTCAATATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTC 871
Db 481 AATATCTCAATATATTCATATTCATATTCATATTCATATTCATATTCATATTCATATTC 540
QY 872 AATGTTTCTGATCTCAAAATGTCGAGCCCAAAATATTAATTGTCGATCATGTA 931
Db 541 AATGTTTCTGATCTCAAAATGTCGAGCCCAAAATATTAATTGTCGATCATGTA 600
QY 932 TA 933
Db 601 TA 602
RESULT 10
LOCUS AF152758 1541 bp mRNA linear PLN 21-MAY-2000
DEFINITION Actinidia chinensis clone CKPgA-3 polygalacturonase A (pGa) mRNA.
ACCESSION AF152758
VERSION AF152758.1 GI:7959982
KEYWORDS
SOURCE
ORGANISM Actinidia chinensis
Actinidia chinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Ericales; Actinidiales; Actinidia.
REFERENCE 1 (bases 1 to 1541)
AUTHORS Wang,Z.Y., Macrae,E.A., Wright,M.A., Bolitho,K.M., Ross,G.S. and
Ackinson,R.G.
TITLE Polygalacturonase gene expression in kiwifruit: relationship to
fruit softening and ethylene production
JOURNAL Plant Mol. Biol. 42 (2), 317-328 (2000)
MEDLINE 20252524
PUBMED 10794531
REFERENCE 2 (bases 1 to 1541)
AUTHORS Ackinson,R.G.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-1999) Plant Development, HortResearch, Private
Bag 92169, Auckland, New Zealand
FEATURES
source Location/Qualifiers
1..1541
/organism="Actinidia chinensis"
/mol_type="mRNA"
/db_xref="taxon:3625"
/clone="CKPGA-3"
/tissue_type="fruit pericarp"
1..1541
/gene="pGa"
1..1392
/gene="pGa"
/gene="pGa"
/codon_start=1
/product="polygalacturonase A"
/protein_id="AAFP71160.1"
/db_xref="GI:7959983"
/translation="MALORHFFQFVITLLIPSFLLGYTSAVHEDPPHDYHLEBYGD
FKAYPSYITITGDNDFGSMHENGIFGLIRKVDYDMRVLDAISKTVVNDVDFGAKDGS

DDTKAFKAMKAVCSSTSSAVLLVPQKNTVREPTSGRCKSLTMQITGTELASDR
 SYSKGRHMLVVDVSQONLRVEGGGIIINGQJLWMONSCITNTKLPCKAPLALTYK
 SKHIVKLNKLENAQOIHSFDCNVAVPASNMTVPSPNPDGHHVGTQNIHIS
 CVIETGDDICISIVSGSKRVNDITCGPBGHSIGSLGNSNASHSDVYNARAKCG
 TTNGVRIKTWOGSGSASNIKFQNMENHNVENPIIIDONKCDPDKCOEBSASVQYKN
 IFYKNTKGTCSANVAITFDCKSRKFPQGIIVLEVDIEIEGGAALKCNVELSEITGV
 VSPHCEBGEBSAS"

Query Match 33.6%; Score 549.2; DB 8; Length 1541;
 Best Local Similarity 70.0%; Pred. No. 1.9e-76;
 Matches 785; Conservative 0; Mismatches 328; Indels 9; Gaps 3;

Db 270 AAAAGTATTAATGTTAGCTTTGAGCTTAAGGCTGATGGAACAAATGATTAATTT 329
 262 AAAAGTATTAATGTTAGCTTTGAGCTTAAGGCTGATGGAACAAATGATTAATTT 318
 QY 330 GCAATTGAGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 389
 Db 319 GCTTTGAGCAAGCTTGAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 378
 QY 390 CCTAAAGCAAGATTAATGTTGAGCAATGATGATGATGATGATGATGATGATGATGAT 449
 Db 379 CCCCAG--AAGAACTATCTTGTAGGCCAATAACATTCAGGCCCATGCAATCTGAT 435
 QY 450 ATTTCAGTAAAGTTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 506
 Db 436 CTCACATGAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
 QY 507 GATAGAGGCTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 566
 Db 496 GATGAGAGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555
 QY 567 ACTATCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
 Db 556 ACCATCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 615
 QY 627 CCATGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 686
 Db 616 CCTGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
 QY 687 AACTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746
 Db 676 AACTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
 QY 747 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 806
 Db 736 CAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
 QY 807 GATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 866
 Db 796 GTCACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
 QY 867 ATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 926
 Db 856 ATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
 QY 927 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 986
 Db 916 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
 QY 987 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1046
 Db 976 GTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1035
 QY 1047 GAGAT 1106
 Db 1036 GGGGAT 1095
 QY 1107 TATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1166
 Db 1096 AACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1155

QY 1167 TCAGCAGCTTCAAGTGAAGAAATGCTGTATGAGATATCAAGGCGACACAGTCAACAAAG 1226
 Db 1156 TCAGCAGCTTCAAGTGAAGAAATGCTGTATGAGATATCAAGGCGACACAGTCAACAAAG 1215
 QY 1227 GTGGCAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286
 Db 1216 GTGGCAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
 QY 1287 ATAAATTTAGTGAAGGGAAGTGAAGAAACCATGAGGCTACGTGCAAAATGCTCATTTT 1346
 Db 1276 GTTGAATCTAGAGTATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1335
 QY 1347 AACATGCTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
 Db 1336 AGTGAAGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
 RESULT 11
 LOCUS AY043233 1595 bp mRNA linear PLN 20-MAY-2003
 DEFINITION Vitis vinifera polygalacturonase mRNA, complete cds.
 ACCESSION AY043233
 VERSION AY043233.1 GI:15081599
 KEYWORDS
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicoties;
 rosids; Vitaceae; Vitis.
 REFERENCE 1 (bases 1 to 1595)
 AUTHORS Nunan,K.J., Davies,C., Robinson,S.P. and Fincher,G.B.
 TITLE Expression patterns of cell wall-modifying enzymes during grape
 berry development
 JOURNAL Planta 214 (2), 257-264 (2001)
 MEDLINE 2158619
 PUBMED 11803390
 REFERENCE 2 (bases 1 to 1595)
 AUTHORS Fincher,G.B.
 TITLE Direct Submision
 JOURNAL Submitted (02-JUL-2001) Dept. Plant Science, University of
 Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia
 REFERENCE 3 (bases 1 to 1595)
 AUTHORS Nunan,K.J., Davies,C. and Robinson,S.P.
 TITLE Direct Submision
 JOURNAL Submitted (02-JUL-2001) Plant Industry, CSIRO, Hartley Grove,
 Adelaide, SA 5064, Australia
 FEATURES
 source location/Qualifiers
 1..1595
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cullivar="Shiraz"
 /db_xref="taxon:29760"
 55..1389
 /codon_start=1
 /product="polygalacturonase Pg1"
 /protein_id="AAK81876.1"
 /db_xref="GI:15081600"
 /translation="MSLORFVPCILVPLISPCVSLQASPPYDLGEASGYDSQAN
 HHHRSTORRETNMPMKRFGSKTGLASLVKMNYYGAGKDGSDAPKARKA
 ACSRSQSLVVPENKNTLKPITFQPCSSITVQIVGVQASTDSASNDTHWLI
 FENVQNLAVQGGTNGNGKTWENSCKNYDLPCKGAPLALFVNCMLAVNLKIQ
 NAOQMIVSEKCVQASGLITTAPENSPNDGIVHSYTKIIOISSVIGTGDICSI
 VSGSQVYNGITCGPHGISIGSLGSDSEAHVSVTVNGATLSGTGVRKTWOG
 GSGSANTIKFQNIWNVNENPIIIDQYCDQSPCKSSASAVOVQVIVQNIKGTSS
 KEAISLDCAKRPPCGIILRIDIKYGGKAKAKAVSNARVYTMGVSPPCA"

ORIGIN
 Query Match 32.1%; Score 525.2; DB 8; Length 1595;
 Best Local Similarity 65.6%; Pred. No. 1e-72;
 Matches 815; Conservative 0; Mismatches 418; Indels 9; Gaps 3;

QY 268 TTTAACTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327

Db 278 TTAATAGTTAAAGTAAATATTATAGAGCTAAAGTGATGG--ANGATGTCACAG 334
Qy 328 TTGATTTGACAGACATGATGATGATGTTTATCTAGAAACACTTGTCAATTTGGG 387
Db 335 AGGATTCAGAAAGCTTGAAGGCGTTGTATCTCCAGAGATCTGTCTA--GTGG 391
Qy 388 TTCCAAAACAGAGATATCTCTTCAGACAAATCACCTTTTCAGTCCATGACAGATTT 447
Db 392 TGCCCAAAAACAGAGATATCTCTTCAGCCAAATTTACATTTCAAGTCTCTGTAAATCA 451
Qy 448 CTATTTTCAGTAAAGATTTTGGATCTTTAGAGCATCTAGTAAATTTCAAGCTAC--A 504
Db 452 GTATTTACAGTGCAGATATATGGAACCTGTCAGGSCATCTACTGATGATGCCCTACAGTA 511
Qy 505 AAGATPAGAGGCTTTGATGCTGTTGTATGTTTCAAAATTTAGTTGTTGGAGAGAG 564
Db 512 ACGATATGACCACTGCTTATCTTTGAAAATGTCMAAATTTAGCGGTTCAAGGTGTG 571
Qy 565 GAACTATCAATGCAATGAGACAAAGATGTGAGCCAAAGTTCTTGGCAAAATTAATTAATCAC 624
Db 572 GCACCATCAATGGAATGGGAAGAGCTGTGGAAAACCTCTTGCAGAGTGAATATGATC 631
Qy 625 TGCCATGACAGGATGACCAACGGCTTAACTCTTGTGAATTTGCAAAATTTGAAAGTGA 684
Db 632 TTCTTTGCAAGGGTGCACCAACGGCTTAACTCTTATACCTGCAAGAACTTGGCAGTGA 691
Qy 685 ATAACTPAAAGATPAAATGACAAAGAAATTCATGAAATTTGAGTCATGCACTATG 744
Db 692 AGAACCTCAAAATCCAAATATGCCAGCAAAATGCAATGATTTCTTTGAAAATAGCGTGGGTG 751
Qy 745 TTGATGCTTCAAAATTTGATGATCAATGCTTCAGCAAAAGGCCAAATCTGATGGAGTCC 804
Db 752 TTCAAGCTCGGGGTCTCCCAATCTGACCCGGGGAACAGTCCCAAGCTGATGAAATCC 811
Qy 805 ATGATCAAAATCTCAATATATTCOAATATCTGATCTATTTATTTGGAACAGGTGATG 864
Db 812 ATGTAGATTAACCAAAATTTATTCAGATTTCAAGCTCGTATTGGAACAGGTGATGACT 871
Qy 865 GTATTTCAATGTTTTCGATCTCAAAATTTGAGGCGCAAAATTTACTGTGGTCCAG 924
Db 872 GTATTTCAATGTTGATGATCCAGAAAGTGCAGATGATGATCACTGTGACCAAG 931
Qy 925 GTCATGTATPAGATTTGGAAGCTTTAGATCTGCAAAATTTAGAACTTATGTCTAATG 984
Db 932 GCCATGGAATCAGCATTTGGAAGCTTGAATCTGAGATTCGAGGCTCATGTTGGAATG 991
Qy 985 TTACTGTAAATGAACCCAAATTTATCGGTGCCGAAAATGAGATTGATCAAGACTTGGC 1044
Db 992 TTACAGTGAACGGGCGCAACGCTTTCGGGAAACCAAAATGAGATTGATCAAGACATGGC 1051
Qy 1045 AGGAGAGATCTGAGCAGCTAGCAAAATTTCTGAATTTGGAATTTGCAAGACCTTA 1104
Db 1052 AGGAGAGGCTGGAAGCGCAAGCAACATCAAAATTTCAAAATTTGATTAATGATTAACGCTGG 1111
Qy 1105 AGTATCCCATATATATAGACCAAAATCTATTTGATCGATGGAACCATGTATACAAACAGT 1164
Db 1112 AGAACCCATATATATATAGCAAAAGTACTGTGACCAAGCAAAACATGACAGTACAGTA 1171
Qy 1165 TTTGAGCAGTTCAAGTGAATAATGTGTATGAGATATCAAGGCGCAAAAGTGCACAA 1224
Db 1172 GCTCGGCGTGTCAAGTGCATAATGTGTGTACAGAAATTAAGGAACAAAGTTCTTGTGA 1231
Qy 1225 AGGAGGCGCATAAATTTGATTTGAGCAGCAAACTTTCATGGAAGGAATTTAATGAGTA 1284
Db 1232 AGGAGGCTATATCGCTTATTTGCAAGTGCAGAGTTTCAATGCAAGGATTTCTGCTGGCAG 1291
Qy 1285 ATATATAATTTAGTAGGGAAGTGAATAACATCAGAGGCTTACGCAAAAATGTCCATT 1344
Db 1292 ACATTTGATATAAAGTTGAGAGAGGAAGGACGCAAGAGCTGTATGACAGCATGCGCAAG 1351
Qy 1345 TTAACAATGCTGAACATGTTACCAACCACTGCATCTCACTAGAAATTTCAAGAGATGAAG 1404
Db 1352 TGACTGTATATGGGAGATGTGTCTCCAAATTTGCGCTGAGAAATGATGGAAGCTAGGATT 1411

Qy 1405 CTCTTTGTATATATTAATTAATTAATGATAGATCTTCATATATATAGCAGATATGATATA 1464
Db 1412 ATTGTTGTGATCTCTGATCAACACTTACATATATATATTTTGTATATATATATATAC 1471
Qy 1465 TCACATTAACAAATCTATATCTATGATTTGATTAATTAATTA 1506
Db 1472 AGTACTGAAAAATATATATGATTTGTGCATATTTGTTATATCA 1513

RESULT 12
TOMPGAAA
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

JOURNAL
COMMENT

FEATURES
source
prim_transcript
CDS

1..7456
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
/note="polygalacturonase mRNA and introns"
join(1479..1757,2416..2547,3327..3491,3696..3716,
4260..4467,4567..4648,5602..5710,6139..6255,6788..7048)
/note="polygalacturonase"
/codon_start=1
/protein_id="AAA34178.1"
/db_xref="GI:170473"
/translation="MVIQNSILLLIIFASSISTCRSNVIDNLFKOVNIIEOEF
AHDFOAYLSTYKSKYILKQITDQKGRSISYKISGLPSSKISPYKRLMIAF
CSSPTPOVQPYPRKXNYLKOITDQKGRSISYKISGLPSSKISPYKRLMIAF
DSVNLVVGGGGTTINGQYWPSSCKINSLPRDPAFTATPNNCKLAKNNLSKN
AQOIHIRESCNTVNASLMINASAKSPNDGVASNTQYQIISDTIIGTDDCISV
SGSONVQATNTTCGPHGISIGSLGSGNSEAYVNTVNEAKIIGAENGVRIKWQSG
SGOASNIKFLNEMQDYKYPILIDQNCDRVEPCIOGSAVQKNNVYENIKGTSATK
VAIKPDSCTNPRGIIIMENINLVGSEKSEBAICKVHFNAHVPHTCSLEISD
EALYNY"
<1479..1757
/note="polygalacturonase"
/number=1
1758..2415
/note="polygalacturonase intron A"
2416..2547
/number=2
2548..3326
/note="polygalacturonase intron B"
3327..3491
3492..3695
/note="polygalacturonase intron C"
3696..3716
/note="polygalacturonase intron D"
4260..4467
4468..4566
/note="polygalacturonase intron E"
4567..4648
/number=6
4649..5601
intron

/note="polygalacturonase intron F (no splice consensus);
putative"
exon 5602..5710
/number=7
intron 5711..6138
/note="polygalacturonase intron G"
exon 6139..6255
intron 6256..6787
/note="polygalacturonase intron H"
exon 6788..>7048
/note="polygalacturonase"
/number=9

ORIGIN

Query Match 28.2%; Score 462; DB 8; Length 7456;
Best local Similarity 100.0%; Pred. No. 5,3e-63;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1161 CAGTTTCAGCACTTCAAGTGAATAATGTGCTATGAGAAATVCAAGGGCACAGTGCA 1220
DB 6785 CAGTTTCAGCACTTCAAGTGAATAATGTGCTATGAGAAATVCAAGGGCACAGTGCA 6844
QY 1221 ACAAAAGTGGCAATAAATTGATGTCAGACAACTTCCAGTGAAGGAATTATATATG 1280
DB 6845 ACAAAAGTGGCAATAAATTGATGTCAGACAACTTCCAGTGAAGGAATTATATATG 6904
QY 1281 GAGATATATAATTAGTAGGGGAAAGTGAACATCAGAGGCTACGTGCAAAAATGTC 1340
DB 6905 GAGATATATAATTAGTAGGGGAAAGTGAACATCAGAGGCTACGTGCAAAAATGTC 6964
QY 1341 CATTTTAACAATGCTGAACATGTTACACCACTGCACTTCACTAGAAAATTCAGAGAT 1400
DB 6965 CATTTTAACAATGCTGAACATGTTACACCACTGCACTTCACTAGAAAATTCAGAGAT 7024
QY 1401 GAAAGCTTTGTAT 1460
DB 7025 GAAAGCTTTGTAT 7084
QY 1461 TATATCAATATAACAAATCTATATCTATGATTTGAATATATATATATATATATAT 1520
DB 7085 TATATCAATATAACAAATCTATATCTATGATTTGAATATATATATATATATATAT 7144
QY 1521 TTGAAGTTTATATAGACTATATATTTCTATTTTCTAGTCAAAAAGTTTGACGATTGT 1580
DB 7145 TTGAAGTTTATATAGACTATATATTTCTATTTTCTAGTCAAAAAGTTTGACGATTGT 7204
QY 1581 ACTTTTAAATGTACAAAATAATATAAATGGTTATTTATATATGA 1622
DB 7205 ACTTTTAAATGTACAAAATAATATAAATGGTTATTTATATATGA 7246

RESULT 13
LEPOLYGA 7456 bp DNA linear PLN 09-MAY-1995
LOCUS
DEFINITION Tomato gene for cell wall degrading polygalacturonase.
ACCESSION X14074
VERSION X14074.1 GI:19305
KEYWORDS polygalacturonase.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 7456)
AUTHORS Bird,C.R., Smith,C.J., Ray,J.A., Moureau P., Bevan,M.W., Bird,A.S.,
Hughes,S., Morris,P.C., Grierson,D. and Schuch,W.
TITLE The tomato polygalacturonase gene and ripening-specific expression
in transgenic plants
JOURNAL Plant Mol. Biol. 11, 651-662 (1988)
FEATURES
source 1..7456
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"

/cultivar="Ailsa Craig"
/db_xref="taxon:4081"
/clone_1b="lambda EMBL3"
repeat_region 877..904
/note="direct repeat 1"
repeat_region 908..935
/note="direct repeat 1"
TATA_signal 1326..1335
CAAT_signal 1342..1347
CAAT_signal 1400..1403
precursor_RNA 1411..7244
/note="primary transcript"
exon 1411..1757
/number=1
TATA_signal 1431..1435
join(1479..1757,2416..2547,3327..3491,3696..3716,
4260..4467,4567..4648,5602..5710,6139..6255,6788..7048)
/codon_start=1
/product="polygalacturonase"
/protein_id="CAA32235.1"
/db_xref="GI:295813"
/db_xref="GOA:P05117"
/db_xref="SWISS-PROT:P05117"
/translation="MYIQNSILILITFASISITCRSNVIDNLFKQYDNLLEQF
AIDFQAYLSLTKNIESNNIDKVDKNGIKVINVLFGAKGDKTIDNTAFEDAMBA
CSSRTPVQFVPRKKNRYLKOITFSGPCRSSISVKLFGLSLEASKISDYKDRLLMAF
DSVQNLVVGSGTINNGOVWPPSSCKPKSKLPCRDAPALTFMNCNKLKYNLKSXN
AQQIHKFESCTNVVANSIMINASAKSPNTDGVVSNVTQYIOISDTLIGGDCISIV
SSQSNVQATNITCGRGHGISIGSISGNSAEVYNSNTVNEAKITGAMGPRITWGG
SSQASIKRFLNEMQDYKPIITIDNYCPRVCEICQFSAVQKYNVYENIKTSATK
VAIKDCSTNPFCEGILMENINDVSGSKPSEATCKNVHFNNAEHTPHCTLSISD
EALILVNY"

intron 1758..2415
/number=1
exon 2416..2547
intron 2548..3326
/number=2
exon 3327..3491
intron 3492..3695
/number=3
exon 3696..3716
intron 3717..4259
/number=4
exon 4260..4467
intron 4468..4566
/number=5
exon 4567..4648
intron 4649..5601
/number=6
exon 5602..5710
intron 5711..6138
/number=7
exon 6139..6255
intron 6256..6787
/number=8
exon 6788..7244
/number=9

ORIGIN

polya_site 7244
Query Match 28.2%; Score 462; DB 8; Length 7456;
Best local Similarity 100.0%; Pred. No. 5,3e-63;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1161 CAGTTTCAGCACTTCAAGTGAATAATGTGCTATGAGAAATVCAAGGGCACAGTGCA 1220

Db	6785	CAGTTTCAGCAGTTCAAGTGAAAAAAGTGGTATGAGAAATATCAAGGGCAAGTGCA	6844
Qy	1221	ACAAAGGTGGCCATAAATTTTGTTGTCAGCAACAACCTTTCATGTGAAGCAATTATATG	1280
Db	6845	ACAAAGGTGGCCATAAATTTTGTTGTCAGCAACAACCTTTCATGTGAAGCAATTATATG	6904
Qy	1281	GAGAAATATTAATTTAGTAGGGGAAAGTGAAACCATCAGAGGGCTAGTGCAAAAATGTC	1340
Db	6905	GAGAAATATTAATTTAGTAGGGGAAAGTGAAACCATCAGAGGGCTAGTGCAAAAATGTC	6964
Qy	1341	CATTTTAACAATGCTGAACATGTTACACCACTGCACCTTCATAGAAATTTGAGAGAT	1400
Db	6965	CATTTTAACAATGCTGAACATGTTACACCACTGCACCTTCATAGAAATTTGAGAGAT	7024
Qy	1401	GAACCTCTTTGTATATATTAATTTATATTAATGATCTTCATATATGACAAATATGA	1460
Db	7025	GAACCTCTTTGTATATATTAATTTATATTAATGATCTTCATATATGACAAATATGA	7084
Qy	1461	TATATCACAATAACAATCTATATCTATGATGAAATATATTTATTAATATGACGA	1520
Db	7085	TATATCACAATAACAATCTATATCTATGATGAAATATATTTATTAATATGACGA	7144
Qy	1521	TTGAAGTTTATATAGACTACTATGATTTCTATTTTCTAGTCAAAAAGTTTGACGATGT	1580
Db	7145	TTGAAGTTTATATAGACTACTATGATTTCTATTTTCTAGTCAAAAAGTTTGACGATGT	7204
Qy	1581	ACCTTTTAATGACAAAATTAATTAATGGTATTTTAATATGA	1622
Db	7205	ACCTTTTAATGACAAAATTAATTAATGGTATTTTAATATGA	7246

RESULT 14				
AKI17942		1600 bp	mRNA	linear
LOCUS	AKI17942			PLN 06-DEC-2002
DEFINITION	Arabidopsis thaliana At3g57510 mRNA for putative endo-polygalacturonase, complete cds, clone: RAFL19-13-H11.			
ACCESSION	AKI17942			
VERSION	AKI17942.1 GI:26450945			
WORDS	Flr cDNA; CAP trapper.			
SOURCE	Arabidopsis thaliana ('thale cress')			
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1			
REFERENCE	Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J., Nakajima,M., Enju,A., Kamiya,A., Nanusaka,M., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K. Arabidopsis thaliana full-length cDNA Published Only in Database (2002) 2 (bases 1 to 1600)			
TITLE	Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J., Nakajima,M., Enju,A., Kamiya,A., Nanusaka,M., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K. Direct Submission			
JOURNAL	Submitted (25-NOV-2002) Motooki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:mseki@gsc.riken.go.jp), URL:http://pfjweb.gsc.riken.go.jp, Tel:81-45-503-9625, Fax:81-45-503-9586)			
COMMENT	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:107-120; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FL-C-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified Bluescript vector. Please visit our web site (http://pfjweb.gsc.riken.go.jp/) for further details.			
FEATURES	location/Qualifiers			
SOURCE	1..1600 /organism="Arabidopsis thaliana" /mol_type="mRNA" /cultivar="Columbia"			

```

ORIGIN
Query Match
Best Local Similarity 62.3%; Score 449.2; DB 8; Length 1600;
Matches 788; Conservative 0; Mismatches 443; Indels 33; Gaps 4;

gene
CDS
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL19-13-H11"
/note="common name: thale cress"
1..1600
/gene="At3g57510/T8H10_110"
100..1395
/gene="At3g57510/T8H10_110"
/codon_start=1
/product="putative endo-polygalacturonase"
/protein_id="BAC42580.1"
/db_xref="GI:26450946"
/translacion="MARCCRHIAVFLCVLMLSTCKALSNVDDGCHDESPFSDSLKANDVYLSISDEETLEASTVSYSNFGAKGDKTDDTQAFKAKMKACSTGVTTFLVPEKTYVILKSTRFRFGPCKSLNFPILGLSASTSRSDYKDKNHHMLIEDVNNLSIDGSGTGIINGNGKTMWONSCKIDKSPKCECTAPATLTYLNKLNWKNLRYENAQOIQISIECKENGVENSVEITAPGDSNPDEGJHITWTLIRVNSPIDIGDQCSILEDSTONILQIFLNGCPGHGITSIGSLDEKNSKAYVGINVDGAKFSESNDGVRITKYQSGSGTAKKIKKONIMENWKNPILIIDQYCDPKCKEDDESAVQVKNYVKNISGISATDVALLTNCSEKTPCGIVLENYKIKGISTASCNNANVKNQGVTSFKS"

```

Query Match	27.5%	Score 449.2	DB 8	Length 1600
Best Local Similarity	62.3%	Pred. No. 7.7e-61		
Matches 788	Conservative 0	Mismatches 443	Indels 33	Gaps 4
QY	277	TTAATGACTTAGCTTGGAGCTTAAGGATGATGAAAAACATATGATATATATTCATCTG	336	
Db	302	TTAGCTTTCAAACTTTCGGAGCCAAAGAGATGAAAAACATGATGATATCTCAGGCTTCA	361	
QY	337	AGCAAGCATGGAAATGAAGCATGTTCATCTAGACAACCTGTTCAAATTTGTGTTCCATAA	396	
Db	362	AGAAAGCATGGAAAGAAAGCATGTTCACAAATATGAGTACTACTTCTTGCTTCCATAAG	421	
QY	397	ACAAGATTTTCTTCTTCAGGAATTCACCTTTGAGTCCAATGAGATCTCTTAATTCAG	456	
Db	422	GAAAGACTTATCTCTTAACTCTAAGCTCTAAGCTTTAGAGCCCATGCAAAATCTTACGTA	481	
QY	457	TAAAGATTTTGGATCTCTTAGAAGCATCTAGTAAATTTTGACACTACAAAGATGAAGGC	516	
Db	482	TTTCAGATCTTAGAGCACTTTATACAGCATCTAGAAAGCTTCAGATTAACAAAGCAAAAAC	541	
QY	517	TTTGGATTTGCTTTTGATATGTCTCAAAATTTAGTTGTTGAGAGG-----AGGA	570	
Db	542	ATTGGCTTATCTTTAGAGAGCTTAAACATCTATCAATCGACGCTGCGCAAGGAAATTA	601	
QY	571	TCATGGCCATGACACAGTATGTGGGCCAGTTCTTGCAAAATTAATTAATCAGTCGCAT	630	
Db	602	TTAATGGCAACGGAAAAACCTGGTGGAGAACTATGCAAAATGACAAACTTAAGCCAT	661	
QY	631	GCAGGGATGCAACCAAGGCTTAACTCTTCGSAATTCGAAAAATTTGAAAGTGAATATC	690	
Db	662	GCACAAAAGCTCCAGAGGCTCTTACTTTATACATTTTAAAGAAATTTGATGTGAAGATC	721	
QY	691	TAAAGATTAATAATGCAACAACAATTCATATCAAAATTTGACTATGACCTATATGTTGAG	750	
Db	722	TGAGGGTGAAAAAATGCCAGCAGATTCAGATTCATTTGAGAAATGCAACAAAGTTGAAG	781	
QY	751	CTTCAATTTTGATGATCAATGCTTCACACAAAGAGCCCAATATCAGATGAGAGTCATGAT	810	
Db	782	TTAGTAAATGTGAATCACTGCTCCGGAGCGATATCCACACAGATGATATCATATCA	841	
QY	811	CAAAATCTCAATATTAATTCAAATATCTGATATCTATTTTGGAAACAGGTGATGTTGATTT	870	
Db	842	CTATATCTCAAAACATTCGAGTCTCCAACTCAGATATCGGAACAGGTGATGATTTGATAT	901	
QY	871	CAATTTGTTTCGATCTCAAAATGTGCAAGCCACAATATTAATCTTGTGCTCAGSTCATG	930	
Db	902	CCATTTGAGCATGGAACGCAAAATCTTCAAAATCTTTGATTTAACTTGAGGCCGCCGCTACG	961	
QY	931	GTATAACTATTTGGAAGCTTAGAGATCTGGAATTCAGAAATTCAGAAAGCTATGTCTATATG	990	
Db	962	GGATCACACATTTGGAGCTTGGGGGAGACAGCAATTCGAAAGCTTATGTCTCGGAATTAATG	1021	

QY 991 TAAATGAAGCAAAATTTATCGTCCGAAATGAGTTAGATCAAGACTTGGCAGGAG 1050
| | | | |
Db 1022 TGGATGATGTAAGTTCTCTGAGAGTGCATGAGATTAGATTAAAGCTTATCAGGAG 1081
| | | | |
QY 1051 GATCTGACAGCTAGCAACATCAATTTCTGAAATGGGAAATGCAAGCGTTAACTATC 1110
| | | | |
Db 1082 GATCAGGAAGCTGCAGAAACATTAAATTCAAAATATTCGAAATGGAACGTCAGAAATC 1141
| | | | |
QY 1111 CCATTAATTAAGACCAAACTATTGATGAGTTGAAACCATGATATACAGCTTTGAG 1170
| | | | |
Db 1142 CCATCATTAATCGACGAGCTACTGCGA---CAAGCAAAATGCGAAGACAGAGTCCG 1198
| | | | |
QY 1171 CAGTTCAAGTGAATAATGTCGTATGAGAAATTCAGGGCACAAGTCAACAAGGTGG 1230
| | | | |
Db 1199 CAGTGCAGATGAAAACGTTGTGTACAAAGACATATCTGTACGAGCGCTACGAGTGTGG 1258
| | | | |
QY 1231 CCATTAATTTGATTTGACAGACAAACTTCCATGTGAAGAAATTAATGAGAAATATA 1290
| | | | |
Db 1259 CCATTAACGTTGAATTCGACCGAGAACTATCATGTCAAGGATTTGCTTGAGAACGTA 1318
| | | | |
QY 1291 ATTTAGTGAAGGGAAGTGGAAAACCATCAGAGGCTACGTCAAAAATGTCATTTTACA 1350
| | | | |
Db 1319 AATTAAGAGGAGAA-----CAGCTTCTTGCAAAATGCCAATGTTAAA 1363
| | | | |
QY 1351 ATGCTGAACATTTACACGACCTGCACTTCACTTGAAATTTCAAGATGAAGCTCTT 1410
| | | | |
Db 1364 ATCAAGGACCGCTTCTCTTAATGCTCTTAACGACTAATTAATGTAATGCA----- 1416
| | | | |
QY 1411 TGTATATTAATTTAATTTACTATGATCTTCATATATATAGCAATATGATATATACAA 1470
| | | | |
Db 1417 --CATACATATTTACATGATGATGATGATATTTATATATGATGATATGTTACTACAT 1474
| | | | |
QY 1471 TAAACAAATCTATATCTATGTAATGAATATATTTAATATGTAAGGATGAAGTTT 1530
| | | | |
Db 1475 GCATGCTCTTCTTAATACATGTAAGTAAGTATATGCAAAAATAGTCGAAAGTTTGTTA 1534
| | | | |
QY 1531 AATA 1534
| | | | |
Db 1535 CATA 1538
| | | | |
RESULT 15
AX412563
LOCUS AX412563 1296 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 327 from Patent WO0222675.
ACCESSION AX412563
VERSION AX412563.1 GI:21445021
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
REFERENCE
AUTHORS Glazebrook, J., Wang, X., Dangl, J. L., Eulgem, T. and Zhu, T.
TITLE Plant genes, the expression of which are altered by pathogen
infection
JOURNAL Patent: WO 0222675-A 327 21-MAR-2002;
SYNGENTA PARTICIPATIONS AG (CH); UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US); Glazebrook, Jan (US); Wang, Kun (US); Dangl,
Jeffrey L. (US); Eulgem, Thomas (US)
FEATURES
Source Location/Qualifiers
1..1296
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 27.2%; Score 445.2; DB 6; Length 1296;
Best Local Similarity 64.7%; Pred. No. 3.4e-60;
Matches 720; Conservative 0; Mismatches 368; Indels 24; Gaps 3;

QY 277 TTAATGTAAGTCTTGGAGCTTAAGGATGAGAAAAATATGATTAATTTGCAATTTG 336
| | | | |
Db 203 TTAGTGTTCCTCAACTTCGAGCCAAAGAGATGAAAAATGATATGATATCTCAGGCTTCA 262
| | | | |
QY 337 AGCAGCATGGAATGAGAGATGTTCACTTGAACACCTGTTCATTTGGTTCCTAAA 396
| | | | |
Db 263 AGAAAGCAATGGAAGAAAGATGTTCAACAAATGAGTTACTTCTTGTTCTCTAAAG 322
| | | | |
QY 397 ACAAGATATATCTTCTCAGCAAAATCACTTTTCAGTCCATGAGCAATCTTCAATTTGAG 456
| | | | |
Db 323 GAAAGCTTATCTCTTAAGTCTACTCGATTTAGAGGCCCAATGCAAAATCCTTAACGTAAC 382
| | | | |
QY 457 TAAAGATTTTGGATCTTGAAGATCTAGTAAATTTCAAGCTACAAAGATAGAAAGG 516
| | | | |
Db 383 TTCAGTCTTAGGCACTTATATCAGATCTAGCAAAAGCTTCAATTAACAAAGCAAAAGC 442
| | | | |
QY 517 TTTGATGCTTTTGAATAGTGTCAAAATTTAGTTGTTGAGAGAG-----GGAAGTA 570
| | | | |
Db 443 ATTGCTTAATCTTAAGAGAGCTTAACATCTATCAATCGACGCTGCTCGAGGGAATTA 502
| | | | |
QY 571 TCAATGGCAATGGAACAATATGTTGGCCAGTTCTTGCAAAATTAATTAATCACTGCCAT 630
| | | | |
Db 503 TTAATGGCAACGGAATAACCTGTGTGCAAGACTCAATGCAAAATCGCAAAATCTTAACCCAT 562
| | | | |
QY 631 GCAGGATGCAACCAAGGCTTAACCTTCTGGAATTTGCAAAAATTTGAAAGTGAATATATC 690
| | | | |
Db 563 GCACAAAAGCTCCAAAGGCTCTTAATTAATTAAGATTTGAATGTAAGAAATC 622
| | | | |
QY 691 TAAAGATTAATAATGCAACAATTCATATCAATTAATTTGATGATGACTAATATGTTGAG 750
| | | | |
Db 623 TGAGGATGAATAATCGAGAGATTCAGATTCATTTGATGAGAAATGCAAAAGTTGAAG 682
| | | | |
QY 751 CTTCAATTTGATGATGATGCTTCAAGAAAGAGCCAAATCTGATGAGTCCATGAT 810
| | | | |
Db 683 TTAGTAATGTTGAATGATCTGCTCCGGGATGATCCAAACAGATGATATCCATATCA 742
| | | | |
QY 811 CAATATCTCAATATATCAAAATATCTGATATCTATTTAGGAACAGGTGATGATTT 870
| | | | |
Db 743 CTAATACTCAAAACATTCGAGTCTCCACTGAGATATGGAACAGGTGATGATTT 802
| | | | |
QY 871 CAATTTGTTTGGATCTCAAAATGTGCAAGCCACAATAATTACTTTGGTCCAGGTCAATG 930
| | | | |
Db 803 CCATTTGAGATGGAACGCAAAATCTTCAAAATCTTTGATTAATCTTGCGGCCCGGTCAG 862
| | | | |
QY 931 GTATTAAGTATGGAAGCTTGAATCTGGAATTTCAAGAGCTTATGTTATGTTATCAG 990
| | | | |
Db 863 GGATGCAATGGAGCTTGAGGCGACGACAAATTCGAAAGCTTATGTTCTGGGAATTAATG 922
| | | | |
QY 991 TAAATGAAGCCAAATTTATCGGTGCCGAATATGAGTTAGATCAAGACTTGGCAGGAG 1050
| | | | |
Db 923 TGGATGATGCTAAGTTCTCTGAGAGTGCATATGAGATTAGATTAAAGCTTATCAAGGAG 982
| | | | |
QY 1051 GATCTGCAAGCTTACCAACTCAAAATTTCTGAATGTGAAAATGCAAGCGTTAAAGTATC 1110
| | | | |
Db 983 GATCAGGAACCTGCAGAGCAATTAATTTCAAAATATTCGATGAGAAAACGTCAGAAATC 1042
| | | | |
QY 1111 CCATTAATATAGCAAAACATTTGATGATGAGTTGAAACCTGATATCAAGCTTTGAG 1170
| | | | |
Db 1043 CCATCATATATGACCGAGCTACTGCGA---CAAGCAAAATGCGAAGACCAAGAGTCCG 1099
| | | | |
QY 1171 CAGTTCAAGTGAATAATGTCGTATGAGAAATATCAAGGGCACAAGTCAACAAGAGTGG 1230
| | | | |
Db 1100 CAGTGCAGATGAAAACGTTGTGTACAAAGACATATCTGATGAGCCCTAGAGATGTGG 1159
| | | | |
QY 1231 CCATTAATTTGATTTGACGACCAACTTCCATGTGAAAGATTAATTAATGAGAAATATA 1290
| | | | |
Db 1160 CGATTAAGTTGAATTTGACGAGAGATATCATGTCAAGGATTTGTGTTAGAAACGTGA 1219
| | | | |
QY 1291 ATTTAGTGAAGGGAAGTGAATAACCATCAGAGGCTAGTGAATAATCTCCATTTTACA 1350
| | | | |
Db 1220 AATTAAGAGAGAA-----CAGCTTCTTGCAAAAATGCAATGTTTAAA 1264
| | | | |
QY 1351 ATGCTGAACATGTTACACCACTGCACTTCA 1382
| | | | |

Tue Jul 27 08:55:39 2004

us-10-018-604-1.rge

Page 16

Db 1265 ATCAAGCGACCGTTTCTCTAAATGCTTTAA 1296

Search completed: July 27, 2004, 01:30:44
Job time : 6486 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:00:07 ; Search time 94.9229 Seconds

(without alignment)
1360.306 Million cell updates/sec

Title: US-10-018-604-2

Perfect score: 2390

Sequence: 1 MWIGNSILLILITPASSIS.....VTPHCTSLSEIDELALVNY 457

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2390	100.0	457	1	AAP80299 Polygalac
2	2390	100.0	457	1	AAP94619 Polygalac
3	2390	100.0	457	2	AAR32107 Polygalac
4	2390	100.0	457	4	AAB48338 Tomato po
5	2338.5	97.8	456	6	ABR43936 Tomato en
6	1185	49.6	438	5	ABR92243 Herbicida
7	1127	47.2	431	2	AAW98178 Anthr-sp
8	1127	47.2	431	3	AAW41324 Arabidops
9	1127	47.2	431	5	ABR92704 Arabidops
10	1127	47.2	463	3	AAG41323 Arabidops
11	1119.5	46.8	415	3	AAG41325 Arabidops
12	1115.5	46.7	426	5	ABR92075 Herbicida
13	1103	46.2	433	2	AAW04268 Olieed r
14	1103	46.2	433	2	AAW42649 Brasica
15	1056.5	44.2	404	2	AAR48674 Putative
16	925	38.7	452	4	AAE00422 Tomato po
17	888.5	37.2	441	5	ABR91673 Herbicida
18	845	35.4	436	3	AAG52347 Arabidops
19	845	35.4	459	5	ABR91108 Arabidops
20	845	35.4	459	5	ABR91108 Arabidops
21	840.5	35.2	468	5	ABR91537 Herbicida
22	796.5	33.3	514	2	AAW42122 Japanese
23	791.5	33.1	514	2	AAW04346 Chamaecyp
24	787.5	32.9	434	5	ABR91377 Herbicida
25	787.5	32.9	514	2	AAR74333 Japanese

26	787.5	32.9	514	2	AAR81586
27	787.5	32.9	514	2	AAZ25667
28	787.5	32.9	514	7	ADC34915
29	787	32.9	514	2	AAR53690
30	787	32.9	514	2	AAR69792
31	787	32.9	514	2	AAR93599
32	787	32.9	514	2	AAZ25666
33	787	32.9	514	7	ADC44914
34	767	32.1	460	2	AAR69791
35	764	32.0	453	4	AAW51693
36	764	32.0	507	4	AAW51691
37	741.5	31.0	491	5	ABR90812
38	728	30.5	405	5	ABR92113
39	724.5	30.3	398	5	AAE20568
40	718	30.0	435	5	ABR93472
41	718	30.0	482	3	AAW05118
42	715	29.9	392	5	ABR92116
43	709.5	29.7	515	5	ABR90945
44	708.5	29.6	401	3	AAW41452
45	708.5	29.6	422	3	AAW41451

ALIGNMENTS

RESULT 1
AAP80299
ID AAP80299 standard; protein; 457 AA.

AC AAP80299;
DT 25-MAR-2003 (revised)
DT 20-NOV-1990 (first entry)
XX Polygalacturonase.

XX Fruit ripening; polygalacturonase; pectin esterase.
XX Unidentified.
XX EP271988-A.

XX 22-JUN-1988.
XX 06-NOV-1987; 87EP-00309853.
XX 11-NOV-1986; 86GB-00026879.
XX (ICIL) IMPERIAL CHEM IND PLC.
XX (ZENVE) ZENECA LTD.

XX Bridges IG, Schuch WW, Grierson D;
XX WPI; 1988-169271/25.
XX N-PSDB; AAN80487.
XX Recombinant DNA comprising promoter and terminator sequences - useful in
XX plants for altering ripening properties esp. in tomatoes.

XX Disclosure; Page 7; 22pp; English.

XX This polygalacturonase (PG) is encoded by plasmid clone pTOM6 which is
XX used to produce antisense mRNA (with an inverted sequence to that of PG
XX mRNA) which is inserted into a vector used to transform plants which
XX thereafter have altered ripening properties. The inverted sequence and
XX the PG mRNA form a double-stranded structure which inhibits ex- pression
XX of the PG mRNA. See also AAN80488. (Updated on 25-MAR-2003 to correct PA
XX field.)

XX Sequence 457 AA;

Query Match 100.0%; Score 2390; DB 1; Length 457;
Best Local Similarity 100.0%; Pred. No. 7.3e-194;

Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNLLEQFAHDFQAYLSYSKNIE 60
 XX 1 MWIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNLLEQFAHDFQAYLSYSKNIE 60
 Db 1 MWIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNLLEQFAHDFQAYLSYSKNIE 60

QY 61 SNNNDKVDKNGIKVINVLSFGAKGDKTYDNIAPFQANNEACSSRTPVQFVVPKKNY 120
 Db 61 SNNNDKVDKNGIKVINVLSFGAKGDKTYDNIAPFQANNEACSSRTPVQFVVPKKNY 120

QY 121 LKQITFSGPCRSSISVKIFGSLFASAKSIDYKDRRLMIAPDSYQNLVGGGGTINNGQV 180
 Db 121 LKQITFSGPCRSSISVKIFGSLFASAKSIDYKDRRLMIAPDSYQNLVGGGGTINNGQV 180

QY 181 WPPSSCKINKSLPCRDAPALTFWNCNKLKVNLLKSKNAQOIHKEPESCTNVVANSMLIN 240
 Db 181 WPPSSCKINKSLPCRDAPALTFWNCNKLKVNLLKSKNAQOIHKEPESCTNVVANSMLIN 240

QY 241 ASAKSPNTDGVHVSNTQYIQTISPTIIGTDDCISIVSGSNVQATNITCGPHGISIGSL 300
 Db 241 ASAKSPNTDGVHVSNTQYIQTISPTIIGTDDCISIVSGSNVQATNITCGPHGISIGSL 300

QY 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQGGSGQASNIKFLANVEMODVKYPIIIDON 360
 Db 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQGGSGQASNIKFLANVEMODVKYPIIIDON 360

QY 361 YCDRVEPCIQOFSAYOVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGESG 420
 Db 361 YCDRVEPCIQOFSAYOVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGESG 420

QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLNYY 457
 Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLNYY 457

RESULT 2
 AAP94619 standard; protein; 457 AA.

XX AAP94619;
 AC 25-MAR-2003 (revised)
 DT 21-JUN-1990 (first entry)
 DE Polygalacturonase (PG) cDNA gene product.
 KM Polygalacturonase gene; tomato; ds.
 XX Lycopersicon esculentum.
 OS US4801540-A.
 XX 31-JAN-1989.
 PD 02-JAN-1987; 87US-00000201.
 PF 28-MAR-1986; 86US-00845676.
 PR 17-OCT-1986; 86US-00920574.
 XX (CALO) CALGENE INC.
 PA Hiatt WR, Sheehy RE, Shewmaker CK, Kridl JC, Knauft V;
 PI WPI; 1989-053640/07.
 DR N-PSDB; AAN91112.
 XX Tomato polygalacturonase gene - used for modulating expression in plant
 PT cells or directing expression of heterologous peptide(s).
 XX Disclosure; Page; 8pp; English.
 CC cDNA of gene may act as a probe to the genomic sequence, provides a means
 of modulating the production of PG and acts as a source of the transit

CC peptide which may be joined to heterologous peptides directing them to
 CC cell wall. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 457 AA;
 SQ Query Match 100.0%; Score 2390; DB 1; Length 457;
 Best local similarity 100.0%; Pred. No. 7.3e-194;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNLLEQFAHDFQAYLSYSKNIE 60
 Db 1 MWIQRNSILLIIIFASSISTCRSNVIDNLFKQVYDNLLEQFAHDFQAYLSYSKNIE 60

QY 61 SNNNDKVDKNGIKVINVLSFGAKGDKTYDNIAPFQANNEACSSRTPVQFVVPKKNY 120
 Db 61 SNNNDKVDKNGIKVINVLSFGAKGDKTYDNIAPFQANNEACSSRTPVQFVVPKKNY 120

QY 121 LKQITFSGPCRSSISVKIFGSLFASAKSIDYKDRRLMIAPDSYQNLVGGGGTINNGQV 180
 Db 121 LKQITFSGPCRSSISVKIFGSLFASAKSIDYKDRRLMIAPDSYQNLVGGGGTINNGQV 180

QY 181 WPPSSCKINKSLPCRDAPALTFWNCNKLKVNLLKSKNAQOIHKEPESCTNVVANSMLIN 240
 Db 181 WPPSSCKINKSLPCRDAPALTFWNCNKLKVNLLKSKNAQOIHKEPESCTNVVANSMLIN 240

QY 241 ASAKSPNTDGVHVSNTQYIQTISPTIIGTDDCISIVSGSNVQATNITCGPHGISIGSL 300
 Db 241 ASAKSPNTDGVHVSNTQYIQTISPTIIGTDDCISIVSGSNVQATNITCGPHGISIGSL 300

QY 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQGGSGQASNIKFLANVEMODVKYPIIIDON 360
 Db 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQGGSGQASNIKFLANVEMODVKYPIIIDON 360

QY 361 YCDRVEPCIQOFSAYOVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGESG 420
 Db 361 YCDRVEPCIQOFSAYOVKNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGESG 420

QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLNYY 457
 Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLNYY 457

RESULT 3
 AAR32107 standard; protein; 457 AA.

XX AAR32107;
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 05-JUL-1993 (first entry)
 DE Polygalacturonase.
 KM polygalacturonase; PG; pectin esterase; PG; expression regulation;
 KW fruit softening enzymes; flowering plants; fruiting plants;
 XX antisense RNA.
 OS Lycopersicon esculentum.
 XX EP532060-A1
 PD 17-MAR-1993
 PF 06-NOV-1987; 92EP-00117411.
 PR 11-NOV-1986; 86GB-00026879.
 XX (ICIL) IMPERIAL CHEM IND PLC.
 PA (ZENEC) ZENEC LTD.
 XX Bridges IG, Grierson D, Schuch WW;
 XX

DR WPI: 1993-087084/11.
 XX N-PSDB; AAC8453.
 PT Recombinant DNA for flowering and fruiting plants e.g. tomatoes ripening
 PT control - comprises base sequence for transcription contg. inverted
 PT sequence of bases complementary to bases in anti sense ribonucleic acid
 PT encoding softening enzymes, or gene expression regulation.
 XX
 XX Example 12; Fig 1; 20pp; English.
 CC This is the sequence of polygalacturonase from clone pTOM6. The clone was
 CC used to isolate the PG promoter in the construction of a vector encoding
 CC antisense RNA to the PG cDNA and PG gene. This would be useful to
 CC regulate the expression of the fruit softening enzymes in flowering and
 CC fruiting plants. Such antisense RNA would delay fruit softening. (Updated
 CC on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct
 CC PF field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-
 CC AUG-2003 to correct OS field.)
 XX
 SQ Sequence 457 AA;
 Query Match 100.0%; Score 2390; DB 2; Length 457;
 Best Local Similarity 100.0%; Pred. No. 7.3e-194;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WTQNSILLIIIFASSISTCRSNVIDNLFKQYDNLLEOFHADFQAYLSKNI 60
 DB 1 WTQNSILLIIIFASSISTCRSNVIDNLFKQYDNLLEOFHADFQAYLSKNI 60
 QY 61 SNNNDKVDKNGIKYINVLFGAKGDKTYDNIAPFQAMNACSSRTVPQFVPPKRYL 120
 DB 61 SNNNDKVDKNGIKYINVLFGAKGDKTYDNIAPFQAMNACSSRTVPQFVPPKRYL 120
 QY 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRMLTAFPSVQNLVVGSGGTINGNGY 180
 DB 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRMLTAFPSVQNLVVGSGGTINGNGY 180
 QY 181 WWPSSCKINKSLPCRDAPALTALTFWNCNKLKVNNLKSKNAQOIHIKESCTNVVASNLMIN 240
 DB 181 WWPSSCKINKSLPCRDAPALTALTFWNCNKLKVNNLKSKNAQOIHIKESCTNVVASNLMIN 240
 QY 241 ASAKSPNTDGVHVSNTQYIOISDTIIGDPCISIVSGSQNVQATNTTCGPHGISTSL 300
 DB 241 ASAKSPNTDGVHVSNTQYIOISDTIIGDPCISIVSGSQNVQATNTTCGPHGISTSL 300
 QY 301 GSGNSEAVVSVTVNEAKTIGANGVRITKTOGSGGQASNIKFLNVEMQDYKYPITIDON 360
 DB 301 GSGNSEAVVSVTVNEAKTIGANGVRITKTOGSGGQASNIKFLNVEMQDYKYPITIDON 360
 QY 361 YCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGSG 420
 DB 361 YCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGSG 420
 QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLNY 457
 DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLNY 457
 RESULT 4
 ID AAB48338 standard; protein; 457 AA.
 XX AAB48338;
 AC AAB48338;
 XX 20-APR-2001 (first entry)
 DE Tomato polygalacturonase (PG) enzyme.
 XX
 KM Pectin; pectin methyltransferase; PME; polygalacturonase; PG; tomato; pTOM6;
 KM food product; yogurt; milk; fruit juice; whey drink; de-esterification.
 XX
 OS Lycopersicon esculentum.
 XX

PN WO200078982-A1.
 XX 28-DEC-2000.
 PD 15-JUN-2000; 2000WO-1B000865.
 XX 17-JUN-1999; 99GB-00014209.
 XX (DANI-) DANISCO AS.
 XX
 PI Christensen TME, Kreiberg JD;
 XX
 XX WPI: 2001-091573/10.
 DR N-PSDB; AAC8453.
 XX
 PT Modifying pectin, for foodstuffs preparation, involves transforming host
 PT having pectin methyltransferase (PME) and polygalacturonase (PG) activity by
 PT silencing PG activity, to increase PME to PG ratio.
 XX
 XX Disclosure; Fig 1; 78pp; English.
 XX
 CC The invention provides a new method for modifying pectin that involves
 CC providing a host having pectin methyltransferase (PME) activity and
 CC polygalacturonase (PG) activity, transforming the host by silencing PG
 CC activity to provide an increased PME to PG ratio, preparing a PME extract
 CC from the transformed host, and using the PME extract to modify pectin. A
 CC PME modified pectin is useful for foodstuffs preparation, and to impart
 CC an increased functionality to food products such as yogurt, milk/fruit
 CC juice and whey drinks. PME is useful to reduce the number of ester groups
 CC in a pectin in a block-wise manner, and to de-esterify two or more
 CC adjacent galacturonic acid residue of a pectin on at least substantially
 CC all of the pectin chains. The present sequence represents a PG enzyme
 CC encoded by a pTOM6 cDNA
 XX
 SQ Sequence 457 AA;
 Query Match 100.0%; Score 2390; DB 4; Length 457;
 Best Local Similarity 100.0%; Pred. No. 7.3e-194;
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WTQNSILLIIIFASSISTCRSNVIDNLFKQYDNLLEOFHADFQAYLSKNI 60
 DB 1 WTQNSILLIIIFASSISTCRSNVIDNLFKQYDNLLEOFHADFQAYLSKNI 60
 QY 61 SNNNDKVDKNGIKYINVLFGAKGDKTYDNIAPFQAMNACSSRTVPQFVPPKRYL 120
 DB 61 SNNNDKVDKNGIKYINVLFGAKGDKTYDNIAPFQAMNACSSRTVPQFVPPKRYL 120
 QY 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRMLTAFPSVQNLVVGSGGTINGNGY 180
 DB 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRMLTAFPSVQNLVVGSGGTINGNGY 180
 QY 181 WWPSSCKINKSLPCRDAPALTALTFWNCNKLKVNNLKSKNAQOIHIKESCTNVVASNLMIN 240
 DB 181 WWPSSCKINKSLPCRDAPALTALTFWNCNKLKVNNLKSKNAQOIHIKESCTNVVASNLMIN 240
 QY 241 ASAKSPNTDGVHVSNTQYIOISDTIIGDPCISIVSGSQNVQATNTTCGPHGISTSL 300
 DB 241 ASAKSPNTDGVHVSNTQYIOISDTIIGDPCISIVSGSQNVQATNTTCGPHGISTSL 300
 QY 301 GSGNSEAVVSVTVNEAKTIGANGVRITKTOGSGGQASNIKFLNVEMQDYKYPITIDON 360
 DB 301 GSGNSEAVVSVTVNEAKTIGANGVRITKTOGSGGQASNIKFLNVEMQDYKYPITIDON 360
 QY 361 YCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGSG 420
 DB 361 YCDRVEPCIQFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGSG 420
 QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLNY 457
 DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLNY 457

```

RESULT 5
ABR43936
ID ABR43936 standard; protein; 456 AA.
XX
XX ABR43936;
AC
XX
XX 11-AUG-2003 (first entry)
DT
XX
XX Tomato endopolygalacturonase enzyme.
DE
XX
XX Plant; pectin transferase synthase; xyloglucan; polylysine; heparin;
XX
XX anticoagulant; endopolygalacturonase; enzyme; tomato.
XX
XX Lycopersicon esculentum.
OS
XX
XX WO2003017950-A2.
PN
XX
XX 06-MAR-2003.
PD
XX
XX 03-SEP-2002; 2002WO-US028066.
PF
XX
XX 31-AUG-2001; 2001US-0316777P.
PR
XX
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
PA
XX
XX Albersheim P, Djeleneo-Albersheim I, Darvill A;
PI
XX
XX MPI; 2003-421044/39.
DR
XX
XX Formation of ester or amide bond between ester and alcohol or amine
XX
XX involves treating the ester or its acid or salt with plant pectin
PT
XX
XX transferase synthase in the presence of alcohol or amine.
PS
XX
XX Disclosure; Fig 18; 67pp; English.
XX
XX
XX The invention relates to forming an ester or amide bond between monomeric
XX
XX or polymeric ester or its acid or salt and monomeric or polymeric alcohol
XX
XX or amine by treating the ester or its acid or salt with a plant pectin
XX
XX transferase synthase in the presence of alcohol or amine under conditions
XX
XX to form ester or amide bond. The method is useful for forming an ester or
XX
XX amide bond between monomeric or polymeric ester or its acid or salt (e.g.
XX
XX homogalacturonan) and monomeric or polymeric alcohol or amine. It is
XX
XX useful for producing pectin-based polymers e.g., xyloglucan or D- or L-
XX
XX polylysine, useful for the slow release of compounds in the body e.g.,
XX
XX heparin with anticoagulant or other pharmaceutical properties. The
XX
XX present sequence represents an endopolygalacturonase enzyme from tomato
XX
XX
SQ Sequence 456 AA;
Query Match 97.8%; Score 2338.5; DB 6; Length 456;
Best Local Similarity 98.9%; Pred. No. 1,76-189;
Matches 452; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 MVIQRNSILLIIIFASSISITCRSNVIDNLFKQVYDNIIEQFAHDFOAYLSYLSKNIE 60
DB 1 MVIQRNSILLIIIFASSISITCRSNVIDNLFKQVYDNIIEQFAHDFOAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNTAFEQANNEACSSRTPVQFVVPKKNYL 120
DB 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNTAFEQANNEACSSRTPVQFVVPKKNYL 120
QY 121 LKQITFGPCRSSTISYKIFGSLFASSTISDYKDRRLIAFDSVONLWVGSGGTINNGQV 180
DB 121 LKQITFGPCRSSTISYKIFGSLFASSTISDYKDRRLIAFDSVONLWVGSGGTINNGQV 180
QY 121 LKQITFGPCRSSTISYKIFGSLFASSTISDYKDRRLIAFDSVONLWVGSGGTINNGQV 180
DB 121 LKQITFGPCRSSTISYKIFGSLFASSTISDYKDRRLIAFDSVONLWVGSGGTINNGQV 180
QY 181 WPFSSCKINKSLPCRDAPLALTFFNNCKNLKYNLKSNAQOIHKFESCTNVVANSMLIN 240
DB 181 WPFSSCKINKSLPCRDAPLALTFFNNCKNLKYNLKSNAQOIHKFESCTNVVANSMLIN 240
QY 181 WPFSSCKINKSLPCRDAPLALTFFNNCKNLKYNLKSNAQOIHKFESCTNVVANSMLIN 240
DB 181 WPFSSCKINKSLPCRDAPLALTFFNNCKNLKYNLKSNAQOIHKFESCTNVVANSMLIN 240
QY 241 ASAKSPNTDGVHVSNTYIQLSDITITGDDPCISVSGSQNVQNTITTCGPHGHSIGSL 300
DB 241 ASAKSPNTDGVHVSNTYIQLSDITITGDDPCISVSGSQNVQNTITTCGPHGHSIGSL 300
QY 241 ASAKSPNTDGVHVSNTYIQLSDITITGDDPCISVSGSQNVQNTITTCGPHGHSIGSL 300
DB 241 ASAKSPNTDGVHVSNTYIQLSDITITGDDPCISVSGSQNVQNTITTCGPHGHSIGSL 300

```

```

QY 301 GSGNSEAYVSVNTVNEAKITGAENGVRITWQGGSGQASNIKFLANEMQDVXPIIIDON 360
DB 301 GSGNSEAYVSVNTVNEAKITGAENGVRITWQGGSGQASNIKFLANEMQDVXPIIIDON 360
QY 361 YCDRVEPCIQQFSANQVKNVVTENIKGTSATKVAIKFDCSTNFPCCGIIIMENINYGESG 420
DB 361 YCDRVEPCIQQFSANQVKNVVTENIKGTSATKVAIKFDCSTNFPCCGIIIMENINYGESG 420
QY 421 KPSEATCKNVHFNMAEHVTPHCTSLSEISDEALLVNY 457
DB 421 KPSEATCKNVHFNMAEHVTPHCTSLSEISDEALLVNY 457
RESULT 6
ABR92243
ID ABR92243 standard; protein; 438 AA.
XX
XX ABR92243;
AC
XX
XX 31-MAY-2002 (first entry)
DT
XX
XX Herbicidally active polypeptide SEQ ID NO 1454.
DE
XX
XX Herbicidally active polypeptide SEQ ID NO 1454.
XX
XX Herbicidal; plant; agriculture; herbicide.
XX
XX Arabidopsis thaliana.
OS
XX
XX WO200210210-A2.
PN
XX
XX 07-FEB-2002.
PD
XX
XX 28-AUG-2001; 2001WO-EP009892.
PF
XX
XX 28-AUG-2001; 2001WO-EP009892.
PR
XX
XX (FARB) BAYER AG.
PA
XX
XX Tietjen K, Weidler M;
PI
XX
XX MPI; 2002-269010/31.
DR
XX
XX Identifying plant target proteins for herbicidally active compounds,
XX
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX
XX from plant with nucleic acid or amino acid sequences from non-plant
XX
XX organisms.
PS
XX
XX Claim 5; SEQ ID NO 1454; 261pp + Sequence Listing; English.
XX
XX
XX The invention relates to identifying target proteins (ABR90790-ABR94016)
XX
XX for herbicidally active compounds, comprising aligning and comparing
XX
XX CC nucleic acid or amino acid sequences from plant with nucleic acid or
XX
XX CC amino acid sequences from non-plant organisms using suitable search
XX
XX CC parameters, where plant sequences having an E-value greater by a factor
XX
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX
XX CC The polypeptides or nucleic acids encoding them are useful for
XX
XX CC identifying modulators. The identified modulators are useful as
XX
XX herbicides
XX
XX
SQ Sequence 438 AA;
Query Match 49.6%; Score 1185; DB 5; Length 438;
Best Local Similarity 51.9%; Pred. No. 1,26-91;
Matches 230; Conservative 76; Mismatches 121; Indels 16; Gaps 5;
QY 8 ILLIIIFASSISITCRSNV---IDNLFKQ---YDNIIEQFAHDFOAYLSYLSKNIES 61
DB 5 ILLIIIFASSISITCRSNV---IDNLFKQ---YDNIIEQFAHDFOAYLSYLSKNIES 61
QY 62 NNNIDKVDKNGIKVINLVSFGAKGDKTYDNTAFEQANNEACSSRTPVQFVVPKKNYL 121
DB 62 NNNIDKVDKNGIKVINLVSFGAKGDKTYDNTAFEQANNEACSSRTPVQFVVPKKNYL 121
QY 57 KNRGIVAPRRSPRRFNNVTFAKANGND-DSKAPMKAMEAACSTGATYIYAAPKKNRDYML 115
DB 57 KNRGIVAPRRSPRRFNNVTFAKANGND-DSKAPMKAMEAACSTGATYIYAAPKKNRDYML 115
QY 122 KQITFGPCRSSTISYKIFGSLFASSTISDYKDRRLIAFDSVONLWVGSGGTINNGQV 181
DB 122 KQITFGPCRSSTISYKIFGSLFASSTISDYKDRRLIAFDSVONLWVGSGGTINNGQV 181

```

```
Db      116 KAVTFSGPCKSSLIITFYGRITKEMENPSPDKERHMIWFENVNKLREGGGRIDGNGH1W 175
      182 WPSSECKINKSLPCRDAPALTFMNCCKNLKVNLSKSKNAQOIHKFESCTVVASNM1NA 241
      176 WPRSCINPQLPLCLGAPLATVFECCNNLRVSNRLRNAQOMHILTFQDCNKVKALNLMVMS 235
Qy      242 SAKSEPTDGVHVSNTQYIQISDTITIGTDDCISIVSGSQNVQATNTTGGFHGISIGSLG 301
      236 PADSPNTDGIHVSGTGNMIIIDISIVRTGDDCISIVSGSEVRATGITCGFHGISIGSLG 295
Qy      302 SGNSSEAYVSNVTNNAKTIIGAENGVRITKMOGSSGQASNIKFLANVMQDYKPIITIDQNY 361
      296 EDNSEAYVSNVNVNKKATLIGTNGVRIKTWGGGHGMAKNIIFQDIIMKVTNP1IINQDY 355
Db      362 CURVEPCIOQFSAYQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIENINLVGSGK 421
      356 CURVEACPEQKSAVQVSNVLKXKIOGTSRPIAVKFCVSKNIPCRGISMQNVKALVDTQO 415
Qy      422 P-SEATCKNVHFNNAEHVTPHCT 443
      416 DVSKASCSNVKLDTRGNVSP1CT 438
Db
```

RESULT 7

AAM98178 standard; protein; 431 AA.

AAM98178;

05-UUL-1999 (first entry)

Anther-specific ESJ2A protein.

ESJ2A gene; promoter; pollen; anther dehiscence; male sterile;
transgenic plant.

Arabidopsis thaliana.

WO9913089-A1.

18-MAR-1999.

11-SEP-1998; 98WO-GB002752.

11-SEP-1997; 97GB-00019359.

(BIOG-) BIOGEMMA UK LTD.

Roberts JA, Paul W, Craze M;

WPI; 1999-254279/21.

N-PSDB; AAX25011.

Generation of male sterile plants by controlling anther dehiscence.
Disclosure; Fig 3; 34pp; English.

This protein is encoded by the ESJ2A gene (see AAX25011) of Arabidopsis thaliana. The invention relates to the use of the ESJ2A promoter to reduce dehiscence and to create male sterile plants for use in hybrid seed production. The promoter is used to drive expression of a further nucleic acid sequence that results in prevention or reduction of anther dehiscence. For example, expression of the RNase barnase causes cell ablation, while expression of a plant hormone alters the developmental fate of a cell. Plants are produced that have phenotypically normal pollen grains, within phenotypically normal anthers, but in which the anthers do not dehiscence and thus do not release the pollen grains. The system allows the female to be multiplied with the artificial male sterility gene in the homozygous state, since the female plant produces viable pollen. The system is suited to crops which have high seed multiplication, large amounts of pollen, and/or separate male and female inflorescences. These factors allow for easy collection of pollen from

CC non-dehiscing anthers, facile self-pollination and the minimisation of
CC the area of plants that have to be self-pollinated manually. Such an
CC ideal crop is the monocot maize, but the system is also applicable to
CC e.g. wheat, barley, rice, fodder grass, banana, palm, orchid, tulip,
CC lily, melon, cucumber, tomato, pepper and willow. It is also useful in
CC the avoidance or reduction of pollen allergens and may be effective in
CC the control of asthma caused by pollen release

SQ Sequence 431 AA;

Query Match 47.2%; Score 1127; DB 2; Length 431;

Best Local Similarity 49.6%; Pred. No. 9.4e-87; Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

```
Qy      7 SILLIIIPASSISTCRSNVID-----DNLFKQVONLIEQEFADHQAVLSYL 55
      9 APLFLCULMLSLCKALSSNVDDGYGHEDSFPESDILKLNDDVL-----SLISSD 59
Db      56 SKNIESNNNIDRVKNGIKVINLVSFGAKGDKTYDNIAPEDQAMNEACSSRTPEQFVVPK 115
      60 ETLLEAS-----TVSVNFGAKGDKTDDTQAFKXAMKACSTNGVTFELVPK 107
Qy      116 NKNVILLKQITPESGPCRSSISVKKIFGSLSEASKISDYKDRRLMIAFDSVQNLVVGG- 173
      108 GKTYLLKSTFRFGPKCKSLRNFIQILGTLSTKRSYDKDKNHWLILEDVNNLSIDGSGTGI 167
Qy      174 INNGQVWMPSSCKINKSLPCRDAPALTFMNCCKNLKVNLSKSKNAQOIHKFESCTNVV 233
      168 INNGKTMQNSCKIDKSKPCCKAPALTLVNLKLNVRKNAQOIQISIEKCKVYE 227
Qy      234 ASNLMINASAKSPNTDGVHVSNTQYIQISDTITIGTDDCISIVSGSQNVQATNTTGGFH 293
      228 VSNVEITAPDGSBNTDGIHITNTQNRVSNISDGTDDCISIEDGTQNIIP1LTGGPH 287
Db      294 GISIGSLGSGNSEAYVSNVTNNAKTIIGAENGVRITKMOGSSGQASNIKFLANVMQDYK 353
      288 GISISLGDNDKAYVSGINVDGAKFSESNDNGRIKTYOGSGTANIKRQNRMENNVN 347
Qy      354 PIITDQNYCDRVEPCIOQFSAYQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIENI 413
      348 PIITDQNYCDK-DKEPDQSAVQVKNVYKNSGISATVAILTNSEKIPCOGIVLENV 406
Db      414 NLVSGSKPSEATCKNVHFNNAEHVTPHCT 443
      407 KIKG-----GTASCKNANVKNQGTVPKCS 431
```

RESULT 8

AAG41324 standard; protein; 431 AA.

AAG41324;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 51400.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-00301439.

25-FEB-1999; 99US-0121825P.

05-MAR-1999; 99US-0123180P.

09-MAR-1999; 99US-0123548P.

23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127452P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 03-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139763P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154039P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.

PR	08-OCT-1999;	99US-0158322P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161932P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

Query Match	47.2%	Score 1127	DB 3	Length 431
Best Local Similarity	49.6%	Pred. No. 4e-87		
Matches 223	Conservative 71	Mismatches 116	Indels 40	Gaps 6

QY	7	STLLIIIFASSISTCRSNVID-----DNLFKQVYDNIIEQEFADHFOAYLSYL	55
Db	9	AVFLCVLLMLSTICKALSNVDDGYGHEDEGSFESDSLKLNDDVL-----SLISSD	59
QY	56	SKNISSNNNDKVKDKNGIKVINVLSFGAKGDKTYDNTAFEOAMNEACSSRTPVQFVYVK	115
Db	60	ETTLKAS-----TVSSNFGAKGDKTDTDTOAFKKAMKACSTNGVTFELVYK	107
QY	116	MKNVLLKQITFSGPCRSSISVYKIFSLSEASSKISDYKPRMLIAFDSYQNLVVG--GT	173
Db	108	GKTYLLKSTRFRGPKSLRNQILGLTASSTRSRQYKDKMHLIEDVNNLSIDGSGIGI	167
QY	174	INGNQVWMPSSCKINKSLPCEDATALTFFNNCKRLKYNLKSQAQOIHIFESCTNVV	233
Db	168	INGNKTMWQSSCKIDKSPCTKAPALTLYLKNLNVKNLRVKRAQOIOISIECKNVE	227
QY	234	ASNLINASAKSPNTDGVHVSNTQYIQISDITIGTGDDCISIVSGQVQVQATNITCGRGH	293
Db	228	VSNVEITAPGSPENPDGHIHTNTQNIIRSNPSDIGDDCISIEBGTQNLQFLDTLCGRGH	287
QY	294	GISIGSLGSGSEAVYNSVTYNEAKITGAENGVRITKWQSGSGAASNIKFLVNEQDQYK	353
Db	288	GISIGSLGDDNSKAVYSGINVDGAFKFSBMDGVRITKIQGSGGTAKNIKFNIRKENYKN	347
QY	354	PIIDONYCDEAREPTIQOFSAYOVQNNVYENIKGISATKVAIKPFCSTNPFCEGITMENI	413
Db	348	PIIIDDDYCDK-DKDEDESAVQVKNVYKIIISGISATDVALITLNCSEKTPCOQIVLENV	406
QY	414	NLVBESGKPSALTKCNVHNNAEHVTPHCT	443
Db	407	KIKG-----GTASCKNANVKNQGTIVSPKCS	431

RESULT 9	ABB92704	standard; proteoin; 431 AA.
ID	ABB92704	
XX	ABB92704;	

XX	31-MAY-2002 (first entry)
DT	
XX	Herbicideally active polypeptide SEQ ID NO 1915.
DE	
XX	Herbicideal; plant; agriculture; herbicide.
KW	
XX	Arabidopsis thaliana.
OS	
XX	WO200210210-A2.
PN	
XX	07-FEB-2002.
PD	
XX	
PF	28-AUG-2001; 2001WO-EP009892.
XX	
PR	28-AUG-2001; 2001WO-EP009892.
XX	
PA	(FARB) BAYER AG.
XX	
PI	Tietjen K, Weidler M;
XX	
DR	WPI; 2002-269010/31.
XX	
PT	Identifying plant target proteins for herbicideally active compounds,
PT	comprising aligning and comparing nucleic acid or amino acid sequences
PT	from plant with nucleic acid or amino acid sequences from non-plant
PT	organisms.
XX	
PS	Claim 5; SEQ ID NO 1915; 261pp + Sequence listing; English.
XX	
XX	The invention relates to identifying target proteins (ABB90790-ABB94016)
CC	for herbicideally active compounds, comprising aligning and comparing
CC	nucleic acid or amino acid sequences from plant with nucleic acid or
CC	amino acid sequences from non-plant organisms using suitable search
CC	parameters, where plant sequences having an E-value greater by a factor
CC	of 3 than the E-value of most similar non-plant sequences are selected.
CC	The polypeptides or nucleic acids encoding them are useful for
CC	identifying modulators. The identified modulators are useful as
CC	herbicides
XX	
SD	Sequence 431 AA;

```

Query Match      47.2% Score 1127; DB 5; Length 431;
Best Local Similarity 49.6%; Pred. No. 9.4e-87;
Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6

QY      7 SILLIIIFASSISTCRSNVID-----DNLFKQYDNLTLEQFAHDFQAYLSYL 55
      :::::|||||:::|
Db      9 AFVLCVLLMLSLCKALSSNVDDGCGHEDGSPESDILKLNNDVYL-----SLLSSD 59
      :::::|||||:::|

QY      56 SKINSENNNIDKVDKNGIKVINTLSPEAGKDGKTYNDIAPEQANNEACSSRTPOQFVPEK 115
      :::::|||||:::|
Db      60 ETLLEAS-----TVSASVFLAKDQKTDIDTQAEFKAKAKKACSTGVTTFLEVK 107
      :::::|||||:::|

QY      116 NKVLLIQITFSGGCRSSISVKIFGSLFASAKSIDYQDRRLIMAFDSYQNLVWGSG--GT 173
      :::::|||||:::|
Db      108 GKTYLKSTFRFGPKCKSLRNFQILGTISASTKSDYDCKXNMLLLEDVNNLSIOGSGIGI 167
      :::::|||||:::|

QY      174 INGNQVWMPSSCKINKSLPQRDAPFALTFWNCNKLKVNLLKSKNAQOIHKFESCTNVV 233
      :::::|||||:::|
Db      168 INGNKQWTONSSCKIDKSKPCTKAPFALTLYNLKNLKRLRVKNAQOIISIEKNKVE 227
      :::::|||||:::|

QY      234 ASNLINAKSAKSPPTQSVHSNTOYIOISPTIIIGTGDCTISVGSQNVQATNTTCGGGH 293
      :::::|||||:::|
Db      228 VSNVEITAPEDSPPTDGIHTNTQNIRVNSDDIGTGDCTISIEGTQNLQITFDLTCCGGH 287
      :::::|||||:::|

QY      294 GISIGSLGSGNSEAYVSNVTVNEAKIIGANGVYIKTWOGSGGCAASNIKFLANVEMODVXY 353
      :::::|||||:::|
Db      288 GISIGSLGDDNSKAYVSGINVDGAKFESDNGVRILKTYOGSGGTAKKIHKPONIMENAVKN 347
      :::::|||||:::|

QY      354 PIITDONYCDRVEPCIQOFSAYVQKVVYNIKGTSAKVAIKFDCSTNPFCEGIIMENI 413
      :::::|||||:::|
Db      348 PIITDDYCDK--DKCEQESAVQKVVYNIKGTSAKTDVAITLNGCEKPCQGVLENT 406
      :::::|||||:::|

```


QY 414 NLVBSGKRESEATCKNVHFNNAHVPHCT 443
Db 407 KIRG----GTASCKNVNKNQGTVPKCS 431
RESULT 10
AAG41323
ID AAG41323 standard; protein; 463 AA.
AC AAG41323;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 51399.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 18-MAY-1999; 99US-0134370P.
XX 19-MAY-1999; 99US-0134768P.
XX 20-MAY-1999; 99US-0134941P.
XX 21-MAY-1999; 99US-0135124P.
XX 24-MAY-1999; 99US-0135353P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136783P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140685P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143547P.
PR 14-JUL-1999; 99US-0143634P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145097P.
PR 22-JUL-1999; 99US-0145098P.
PR 22-JUL-1999; 99US-014512P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 02-AUG-1999; 99US-0147038P.
PR 03-AUG-1999; 99US-0147039P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.

```
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 47.2%; Score 1127; DB 3; Length 463;

Best Local Similarity 49.6%; Pred. No. 1e-86; Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

```
QY 7 SILLIIIFASSISTGRNVID-----DNLFQVYDNILBQFAHDFQAVLYSL 55
DB 41 AVPLCVLIMSLCKALSNVDGVDGHDGSEFSDSLKLNNDVYL-----SLISSD 91
QY 56 SKIIEENNNDKDKXGKIKIYNVLSFGAKDQKTYDNIAFEQAWNEAGSRTPVQFVPR 115
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
DB 92 ETTLEAS-----TVSVNFGAKGDKTDDTQAEKKAMKACSTNGVTTFLVPR 139
QY 116 NKNYLLKQITFSGPCRSSISVKIFGSLBASKISDYKDRLLMTAFDSVONTLVVGG--GT 173
DB 140 GKTYYLKSITFRPGPCSLRNFQILGLTISASTKSDYDKXKXNMLLELVNNLSTIGSGTGI 199
QY 174 INNGQVWVPSSCKINKSLPCRDPALTATFWNCKNKVNNLKSKNAQOIHKPESCINNV 233
DB 200 INNGKTMWQNSCKIDKSKRCTKAPALTATLYNKNLNVKNLRKNAQOIISTEKNKVE 259
QY 234 ASNLMTNASAKSNPTGTVHVSNTQYIQISPTTITGDDDCISYSGSNVQATNITGCPGH 293
DB 260 VSNVEITAPDGSNPTGHIHTNTQNRVNSDGTGDDCISIDGTQNLQIFDLTCBPH 319
QY 294 GISIGSLGSGNSEFAYSVNTVNEAKTIGANGVRIKTMQSGGQASITKFLNVMODVXY 353
DB 320 GISIGSLGDDNSGAYVSGIIVDDAKSESBNQVRIKTYQSGSGTAKNIKQNTIMEVKN 379
QY 354 PIIIDQNYCDRVBPCIQQFSAVQKVVYENIKGTSATKVAIKFDCSTNPECGIIMNTI 413
DB 380 PIIIDQYCDK-DKCEDQESAQVQKVVYKNISGTSATDVAILTNGSEKYPQCGIVLENV 438
QY 414 NLVGSAGKPSKSEATCKXVHFNNAEHVTPHCT 443
DB 439 KIRG----GTASCKXNAVKNQGVSPKCS 463

RESULT 11
AAG41325
ID AAG41325 standard; protein; 415 AA.
AC AAG41325;
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 51401.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 22-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
```

PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142802P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157553P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 08-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159332P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.

PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-016120P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 46.8%; Score 1119.5; DB 3; Length 415;
 Best Local Similarity 50.6%; Pred. No. 3.8e-86;
 Matches 223; Conservative 68; Mismatches 107; Indels 43; Gaps 7;

19 ISNCR---SNVID-----DNLFKQVYNNILEQEPAMDFOAYLSLKNIESNNN 64
 2 LSLCKALSSNVDDGYGHEDGSFESDSLKLKLNDDVL-----SLISSDETTLEAS-- 50
 65 IDKVDKNGIKVINVLSPGAKGDKTYDNIAPFQAMNEACSSRTPVQFVFNKNNLYLKQI 124
 51 -----TVSVNFGAKGDKTDQTAFFKAMKACSTNGVTTFVLVFKGTYLLKST 100
 125 TREGPCRSSISVAKIFGSLKSSKISDYKRRRLIAPDSYONLVVGG--GTINGGQVW 182
 101 RFRGPKCSLNFQILGTLASASTKRSYKDKNHWLLEDVNNLSIDGSTGIIINGNGKTMW 160
 183 PSSCKINKSLPGDAPFALTFMNCNKLKVNLSKXNAQOIHIKPESTNVVASLIMNAS 242
 161 QNSCKIDKSKPCFKAPALTLTYMKNLVKRNKAQOIQISIEKCNKVEVSNVETAP 220
 243 AKSPNTDGVHNSVTOYIQISDTYIIGTDDCISIVSGSQNVQANTITGPGHGISIGSLG 302
 221 GDSPTNDGHIHTQNRVNSNDIGTDCISIEDGQNIQIPDLTGPGHGISIGSLG 280
 303 GNSGAYVSNVTVEAKTIGABNGVRIKTWGGSGQASNIKFLNVMODVYKPIIIDQNYC 362
 281 DMSKAYVSGINLVGAKSESDNGVRIKTYOGSGGTAKNIKFNIRMEVKNPPIIIDDYC 340
 363 DREVPCTIQPSAVQVKNVYVENIKGTSATKVAIKFDSTNPCEGIMENINLVGESGKP 422
 341 DK-DKEDQESAQVQVKNVKNISGTSATVAITLNCSEKYPQGGIVLENVKIKG----- 394
 423 SEATCKNVHFNNAEHVTPHCT 443
 395 GTASCKNAVKNQGTVPKCS 415

Db

RESULT 12
 ABB92075
 ID ABB92075 standard; protein; 426 AA.
 AC ABB92075;

XX 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1286.

XX Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

XX WO200210210-A2.

PD 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP009892.

XX 28-AUG-2001; 2001WO-EP009892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidner M,
 PI

XX WPI, 2002-269010/31.
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX Claim 5; SEQ ID NO 1286; 261dp + Sequence listing; English.
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

XX Sequence 426 AA;

Query Match 46.7%; Score 1115.5; DB 5; Length 426;
 Best Local Similarity 49.8%; Pred. No. 8.7e-86;
 Matches 222; Conservative 74; Mismatches 117; Indels 33; Gaps 7;

6 NSILLIIIFASSISTCRSNVIDNLEFKQY-----DNLEQEPAMDFOAYLSLKNIE 60
 6 NUTVFLMALMLFMSCCKASRISPNVYDHSYKFKSDSLIKR--EDITGLRSPVRASLR 63
 61 SNNNDKVDKNGIKVINVLSPGAKGDKTYDNIAPFQAMNEACSSRTPVQFVFNKNNLY 120
 64 TPT-----TVSVDFGAKGDKTDQTAFFVAMKACSSNGAVMLVFKGNTYL 112
 121 LKQITSGPCRSSISVAKIFGSLKSSKISDYKRRRLIAPDSYONLVVGG--GTINGG 178
 113 LKSIQLTGPNCSLTVQIFGLSLASQKRSYKDIKMIWIDGVNLSVDGDTGVVGNG 172
 179 QVWPPSSCKINKSLPGDAPFALTFMNCNKLKVNLSKXNAQOIHIKPESTNVVASLIM 238
 173 ETWQNSCKRNKA-----KALTFVNSKSLIYKLVKRNAQOIQISIEKCSNVQVSNV 225
 239 INASAKSPNTDGVHNSVTOYIQISDTYIIGTDDCISIVSGSQNVQANTITGPGHGISIG 298
 226 VTAPADSPNTDGHITQNRVNSNDIGTDCISIESSQNVQINDITGPGHGISIG 285
 299 SLGSGSEAYVSNVTVEAKTIGABNGVRIKTWGGSGQASNIKFLNVMODVYKPIIID 358
 286 SLGSDNSKAVSGVTVDGAKLSTGDNQVRIKTYOGSGGTASNIIFQNIQMDNVKNPPIIID 345
 359 QNYCDREVPCTIQPSAVQVKNVYVENIKGTSATKVAIKFDSTNPCEGIMENINLVGE 418
 346 QDYCDK-SKCTTKRSANQVKNVYRDISGSASNAITFNCSSKNYPCGGIVLDRVNIKG- 403
 419 SGKPSKATCKNVHFNNAEHVTPHCTS 444
 404 -----GKATCTNVANVDKGAVALPQCNS 425

Db

RESULT 13
 AAW04268
 ID AAW04268 standard; protein; 433 AA.
 AC AAW04268;

XX 16-OCT-2003 (revised)

DT 14-DEC-1996 (first entry)

XX Oilseed rape seed pod polygalacturonase.

DE Polygalacturonase; seed pod; dehiscence; shatter; oilseed rape;

KM transgenic plant; antisense.
 XX

OS Brassica napus; cv. Rafal.
 XX WO9630529-A1.
 XX
 PD 03-OCT-1996.
 XX
 XX 29-MAR-1996; 96WO-GB000757.
 XX PR 31-MAR-1995; 95GB-00006684.
 XX
 XX (NICK-) NICKERSON BIOCHEM LTD.
 XX
 XX Roberts JA, Coupe SA, Jenkins ES;
 PI WPI; 1996-455374/45.
 XX DR N-PSDB; AAT33994.
 XX
 PT Control of seed pod dehiscence - using polygalacturonase or nucleic acid
 PT sequences derived from polygalacturonase gene.
 XX
 XX Example 1; Fig 1; 36pp; English.

CC A polygalacturonase (PG) (AA04248) of oilseed rape cv. Rafal seed pods
 CC is useful for controlling dehiscence. Manipulation of the enzyme's
 CC activity can influence the timing of dehiscence. A cDNA clone (AAT33994)
 CC coding for the PG was isolated from a cDNA library of the rape seed pod
 CC dehiscence zone. PG nucleic acids (patric. antisense) can be used to
 CC regulate dehiscence in crop plants. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX

Sequence 433 AA;

Query Match 46.2%; Score 1103; DB 2; Length 433;
 Best Local Similarity 49.8%; Pred. No. 1e-84;
 Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

```

7 SILIIITFASISITCRSNVID----DNLFKQVYDNLDEFADFPQAYISYSKNIES 61
9 AIFLCVLLMLACCQALSSNVDDGYGHEDEGSFET--DSLTKLNDDVLTLSSDRPTTES 66
62 NNNIDKVDKNGIKYINLVSPFGAKGDKTYDNIAPQANNEACSSRTPVQFVPPKKNYLL 121
67 S-----TVSVSNFGAKGDKTDDTQAFKAKKAKCSTNGVTFPLPKGKTYLL 114
122 KQITFSGPCRSISIVKIFGSLASAKISDY-KDRRLTWAPFSVQNLVVGCG--GTINGNG 178
115 KSIKFRGPKCKLRSPQILGTLASATKSDYSNDKXHWLILBDVNNLSIDGSGAGIVDNG 174
179 QVWPPSSCKINKSLPCRDAPALTFWNCNKLKVNLLKSKNAQOIHIKESCTNVVASNLM 238
175 KIMWQNSCKIDKSKPCRTKAPALTLNINLNVKRLRVNNAQOISIEKCNISVDVKVVK 234
239 INASAKSPNTDGVVNSNTQYIQTPTITIGTDDDCISYSGSONVQANTITGCPHGHSIG 298
235 ITAPGDSPTNTGHIHVATKNIRISNSDGTGDDDCISIDGSONVQINDLTGCPHGHSIG 294
299 SLGSGNSEAVVNSVTVAEAKIIGAENGVRKIKTWOGSGQASNIKFLNVEMODVKYPIIID 358
295 SLGDDNSKAYVSGINVDGATLSETDNGVRKIKTYOGSGTAKIKFQNRMDNVKAPIIID 354
359 QNYCDRYEPCIQOFSAYOVKNVYVENIKGTSATKVALKFDGSTNFPCEGIIEMENINLVE 418
355 QNYCDK-DKCEQOESAVOVNNVVRNIGTSTATVAIMFNCSVVKPCOGIVLENNVING- 412
419 SGRSEATCKRVHFNMAEHVTPHC 442
413 -----GKASCCKNVNVDKGTVPKC 432

```

RESULT 14
 AAY42649
 ID AAY42649 standard; protein; 433 AA.
 XX

AC AAY42649;
 XX
 XX 10-JAN-2000 (first entry)
 XX
 XX Brassica napus Sac66 protein putative sequence.
 DE
 XX Signal transduction protein; dehiscence; male sterile plant; D22 gene;
 XX shatter resistance; oilseed rape; Sac66 protein.
 KW
 XX Brassica napus.
 OS
 XX WO9949046-A1.
 XX PN
 XX 30-SEP-1999.
 PD
 XX 22-MAR-1999; 99WO-GB000905.
 XX PR 20-MAR-1998; 98GB-00006113.
 XX
 XX (BIOG-) BIOGEMMA UK LTD.
 PA
 PI Wyatt P, Roberts JA, WhiteJaw C;
 XX
 XX WPI; 1999-580449/49.
 XX DR N-PSDB; AAZ22980.
 XX

PT A nucleic acid encoding a signal transduction protein involved in plant
 PT dehiscence, useful for producing shatter resistant male sterile plants.
 XX
 XX Example 6; Fig 15; 71pp; English.

The invention provides a nucleic acid encoding a signal transduction
 protein involved in the process of dehiscence. The nucleic acids and
 proteins are useful for regulating or controlling dehiscence of a pod or
 an anther in a plant, useful in the production of male sterile plants.
 The methods, etc. may be used in production of shatter resistance or
 shatter-delayed plants such as oilseed rape (Brassica napus). The present
 sequence represents a B. napus Sac66 protein putative sequence

Sequence 433 AA;

Query Match 46.2%; Score 1103; DB 2; Length 433;
 Best Local Similarity 49.8%; Pred. No. 1e-84;
 Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

```

7 SILIIITFASISITCRSNVID----DNLFKQVYDNLDEFADFPQAYISYSKNIES 61
9 AIFLCVLLMLACCQALSSNVDDGYGHEDEGSFET--DSLTKLNDDVLTLSSDRPTTES 66
62 NNNIDKVDKNGIKYINLVSPFGAKGDKTYDNIAPQANNEACSSRTPVQFVPPKKNYLL 121
67 S-----TVSVSNFGAKGDKTDDTQAFKAKKAKCSTNGVTFPLPKGKTYLL 114
122 KQITFSGPCRSISIVKIFGSLASAKISDY-KDRRLTWAPFSVQNLVVGCG--GTINGNG 178
115 KSIKFRGPKCKLRSPQILGTLASATKSDYSNDKXHWLILBDVNNLSIDGSGAGIVDNG 174
179 QVWPPSSCKINKSLPCRDAPALTFWNCNKLKVNLLKSKNAQOIHIKESCTNVVASNLM 238
175 KIMWQNSCKIDKSKPCRTKAPALTLNINLNVKRLRVNNAQOISIEKCNISVDVKVVK 234
239 INASAKSPNTDGVVNSNTQYIQTPTITIGTDDDCISYSGSONVQANTITGCPHGHSIG 298
235 ITAPGDSPTNTGHIHVATKNIRISNSDGTGDDDCISIDGSONVQINDLTGCPHGHSIG 294
299 SLGSGNSEAVVNSVTVAEAKIIGAENGVRKIKTWOGSGQASNIKFLNVEMODVKYPIIID 358
295 SLGDDNSKAYVSGINVDGATLSETDNGVRKIKTYOGSGTAKIKFQNRMDNVKAPIIID 354
359 QNYCDRYEPCIQOFSAYOVKNVYVENIKGTSATKVALKFDGSTNFPCEGIIEMENINLVE 418
355 QNYCDK-DKCEQOESAVOVNNVVRNIGTSTATVAIMFNCSVVKPCOGIVLENNVING- 412

```


This Page Blank (uspto)

QY 1501 TTATTATTAATGACGATGAGTTTAAATAGACTATGTTTCTATTTCTA 1560
DB 1501 TTATTATTAATGACGATGAGTTTAAATAGACTATGTTTCTATTTCTA 1560
QY 1561 GTCAAAAGTTGACGATGAGTTTAAATAGACTATGTTTCTATTTCTA 1620
DB 1561 GTCAAAAGTTGACGATGAGTTTAAATAGACTATGTTTCTATTTCTA 1620
QY 1621 GAAAAAAGTTGAGTTTAAATAGACTATGTTTCTATTTCTA 1636
DB 1621 GAAAAAAGTTGAGTTTAAATAGACTATGTTTCTATTTCTA 1636
RESULT 3
AAN80487
ID AAN80487 standard; DNA; 1637 BP.
AC AAN80487;
XX
XX 25-MAR-2003 (revised)
DT 20-NOV-1990 (first entry)
XX
XX Plasmid clone pTom6.
DE
XX Fruit ripening; polygalacturonase; pectin esterase; ss.
KM
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH 51..1424
FT /*tag= a
FT CDS /label= polygalacturonase
XX
XX EP271988-A.
XX
XX 22-JUN-1988.
XX
XX 06-NOV-1987; 87EP-00309853.
XX
XX 11-NOV-1986; 86GB-00026879.
XX
XX (ICIL) IMPERIAL CHEM IND PLC.
XX (ZENE) ZENECA LTD.
XX
XX Bridges IG, Schuch WW, Grierson D;
XX
XX WPI: 1988-169271/25.
XX DR P-PSDB; AAB80299.
XX
XX Recombinant DNA comprising promoter and terminator sequences - useful in
XX plants for altering ripening properties esp. in tomatoes.
XX
XX Disclosure; Page ?; 22pp; English.
XX
XX This cDNA clone, encoding polygalacturonase (PG), is used to produce
XX antisense mRNA (with an inverted sequence to that of PG mRNA) which is
XX inserted into a vector used to transform plants which thereafter have
XX altered ripening properties. The inverted sequence and the PG mRNA form a
XX double-stranded structure which inhibits expression of the PG mRNA. See
XX also AAN80488. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 1637 BP; 598 A; 232 C; 290 G; 517 T; 0 U; 0 Other;
SQ
Query Match 99.3%; Score 1625; DB 1; Length 1637;
Best Local Similarity 99.9%; Pred. No. 2.1e-275;
Matches 1636; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 AATCTTTTCAATGACGATGAGTTTAAATAGACTATGTTTCTATTTCTA 60
DB 1 AATCTTTTCAATGACGATGAGTTTAAATAGACTATGTTTCTATTTCTA 60
QY 61 AAGGAATAGATGTTTCTTCTCATTAATATTTTGGCTTCATCAATTTCACTTGAGAA 120
DB 61 AAGGAATAGATGTTTCTTCTCATTAATATTTTGGCTTCATCAATTTCACTTGAGAA 120

DB 61 AAGGAATAGATGTTTCTTCTCATTAATATTTTGGCTTCATCAATTTCACTTGAGAA 120
QY 121 GCAATGTTATGATGACAAATTTATTCACAAAGTTTATGTAATTTCTTGAACAGAAAT 180
DB 121 GCAATGTTATGATGACAAATTTATTCACAAAGTTTATGTAATTTCTTGAACAGAAAT 180
QY 181 TTGCTCATGATTTTCAAGCTTATCTTTCTATTTTGGACCAAAATATGTAAGCAAGATA 240
DB 181 TTGCTCATGATTTTCAAGCTTATCTTTCTATTTTGGACCAAAATATGTAAGCAAGATA 240
QY 241 ATATGACAAAGTTGATTAATAATGGGATTAAGGATTAATGATGATGATGATGATGAT 300
DB 241 ATATGACAAAGTTGATTAATAATGGGATTAAGGATTAATGATGATGATGATGATGAT 300
QY 301 AGGATGATGACAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 360
DB 301 AGGATGATGACAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 360
QY 361 CATCTAGAACACCTGTTCAATTTGTGTTCTTAAACAAAGAAATTTCTTCAAGCAAA 420
DB 361 CATCTAGAACACCTGTTCAATTTGTGTTCTTAAACAAAGAAATTTCTTCAAGCAAA 420
QY 421 TCACCTTTTCAAGCTTCATGACGATCTTCTATTTTCAAGATTTTGGATCTTGAAG 480
DB 421 TCACCTTTTCAAGCTTCATGACGATCTTCTATTTTCAAGATTTTGGATCTTGAAG 480
QY 481 CATCTAGTAATAATTTCAAGCTTCATGACGATCTTCTATTTTCAAGATTTTGGATCTT 540
DB 481 CATCTAGTAATAATTTCAAGCTTCATGACGATCTTCTATTTTCAAGATTTTGGATCTT 540
QY 541 AAAATTTTGTGTTGAGAGAGAGAGAAATCAATGGAATGGAATGGAATGGAATGGAAT 600
DB 541 AAAATTTTGTGTTGAGAGAGAGAGAAATCAATGGAATGGAATGGAATGGAATGGAAT 600
QY 601 GTTCTTGAATAATTAATTAATCACTGACGAGGATGACCAAGGCTTTTAACTTCT 660
DB 601 GTTCTTGAATAATTAATTAATCACTGACGAGGATGACCAAGGCTTTTAACTTCT 660
QY 661 GGAATTTGCAAAATTTGAAAGTGAATTAATTAAGAGTAATAATTAAGAGTAATAATTA 720
DB 661 GGAATTTGCAAAATTTGAAAGTGAATTAATTAAGAGTAATAATTAAGAGTAATAATTA 720
QY 721 TCAATTTGAGTCATGACGATTAATGTTGAGCTTCAATTTGATGATGATGATGATGAT 780
DB 721 TCAATTTGAGTCATGACGATTAATGTTGAGCTTCAATTTGATGATGATGATGATGAT 780
QY 781 AGAGCCCAATATCTGATGAGTCATGATCAATTAATCAATTAATCAATTAATCAATTA 840
DB 781 AGAGCCCAATATCTGATGAGTCATGATCAATTAATCAATTAATCAATTAATCAATTA 840
QY 841 CTATTATTGGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 CTATTATTGGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 CCAAAATTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
DB 901 CCAAAATTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
QY 961 ATTCAAGAGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020
DB 961 ATTCAAGAGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020
QY 1021 ATGAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 ATGAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 TGAATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 TGAATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200

QY	1201	ATATCAAGGCGC	CMASTGCAACAAAGGTGGCCATAAAATTTGATGCAAGCAAACTTTC	1267
Db	1201	ATATCAAGGCGC	CMASTGCAACAAAGGTGGCCATAAAATTTGATGCAAGCAAACTTTC	1266
QY	1261	CATGTGMAAGAAATTAATAATGAGAAATAATAATTTGTGGGGAAAGTGGAAAACCATCAG	1320	
Db	1261	CATGTGMAAGAAATTAATAATGAGAAATAATAATTTGTGGGGAAAGTGGAAAACCATCAG	1320	
QY	1321	AGGCTTCGTGCMAAAATGTCCATTTTAAACATGCTGAACATGTTACACACACTGCACCT	1380	
Db	1321	AGGCTTCGTGCMAAAATGTCCATTTTAAACATGCTGAACATGTTACACACACTGCACCT	1380	
QY	1381	CACATGAAATTTCCAGAGGATGAGAGCTTTTGTATAATTTATTAATTTTAACTATGATCT	1440	
Db	1381	CACATGAAATTTCCAGAGGATGAGAGCTTTTGTATAATTTATTAATTTTAACTATGATCT	1440	
QY	1441	TCAATATATAGCAGATATGATATATACAAATTAACAAATCTATATCTATGATGAATA	1500	
Db	1441	TCAATATATAGCAGATATGATATATACAAATTAACAAATCTATATCTATGATGAATA	1500	
QY	1501	TTATATATTAATAATGACGATTTGAAGTTTAAATGAAGCTACTATGATTTCTATTTCTA	1560	
Db	1501	TTATATATTAATAATGACGATTTGAAGTTTAAATGAAGCTACTATGATTTCTATTTCTA	1560	
QY	1561	GTCAAAG-ATTGACGATTTGACTTTTAAATGACAAAAAATAATAAATGGTTATTATA	1619	
Db	1561	GTCAAAGATTGACGATTTGACTTTTAAATGACAAAAAATAATAAATGGTTATTATA	1620	
QY	1620	TGAAAAAATAAAAAAA 1636		
Db	1621	TGAAAAAATAAAAAAA 1637		

PS	Disclosure; Page ?; 8pp; English
XX	
PT	cells or directing expression of heterologous peptide(s).
DR	P-PSDB; AAP94619.
XX	
XX	
PA	(CALJ) CALGENE INC.
XX	
XX	
PI	Hiatt WR, Sheehy RE, Shewmaker CK, Kridl JC, Knaut V;
XX	
XX	WPI; 1989-053640/07.
XX	
PR	17-OCT-1986; 86US-00920574.
XX	
XX	
PF	02-JAN-1987; 87US-00000201.
XX	
PD	31-JAN-1989.
XX	
PN	US4801540-A.
XX	
FT	/tag= a
XX	
FT	CDS
XX	
OS	Lycopersicon esculentum.
XX	
XX	
KW	Polysaccharuronase gene; tomato; ds.
DE	Polysaccharuronase (PG) gene cDNA.
XX	
DT	21-JUN-1990 (first entry)
XX	
DT	25-MAR-2003 (revised)
AC	AAN91112;
XX	
ID	AAN91112 standard; DNA; 1624 BP.
RESULT 4	

XX	cDNA of gene may act as a probe to the genomic sequence, provides a means
CC	of mobilizing the production of PG and acts as a source of the transit
CC	peptide which may be joined to heterologous peptides directing them to
CC	cell wall. (updated on 25-MAR-2003 to correct PA field.)
XX	
SQ	Sequence 1624 BP; 586 A; 231 C; 294 G; 513 T; 0 U; 0 Other;
	Query Match 98.5%; Score 1611.6; DB 1; Length 1624;
	Best Local Similarity 99.8%; Prid. No. 4.6e-273;
	Matches 1614; Conservative 0; Mismatches 4; Indels 0; Gaps 0
QY	8 TTCAATGACAAAGTTTAAAACCATTGCCATTAATACATATATCATGGTTATCCAAAGGA 67
DB	4 TTCAATGACAAGTTTAAAAAACCATACCATATACATATATCATCGTTATCCAAAGGA 63
QY	68 TAGTATTCCTTCATATATATTTTGCTTCATCAATTTTAACTTTGAGAAGCAATGT 127
DB	64 TAGATATTCCTTCATATATATTTTGCCTCATAATTCACCTTTGAGAAGCAATGT 123
QY	128 TATTGATGACAAATTATTCAAACAAGTTATGATTAATCTTGAACAAGATTTGGCTCA 187
DB	124 TATTGATGACAAATTATTCAAACAAGTTATGATTAATCTTGAACAAGATTTGGCTCA 183
QY	188 TGATTTCAAAGCTTATCTTTCTTATTTGACCAAAAATTTGAAGCAACAATATATTGA 247
DB	184 TGATTTCAAAGCTTATCTTTCTTATTTGACCAAAAATTTGAAACCAACAATATATTGA 243
QY	248 CAAGGTGATTAATAATGGGATTTAAGSTGATTAATGACTTGGCTTTGAGCTAAGGTGA 307
DB	244 CAAGGTGATTAATAATGGGATTTAAGSTGATTAATGACTTGGAGCTTAAGGGTGA 303
QY	308 TGGAAAAACATATGATTAATTTGCACTTTGAGCAAGCATGGAATGAGATGTCATCTAG 367
DB	304 TGGAAAAACATATGATTAATTTGCACTTTGAGCAAGCATGGAATGAGATGTCATCTAG 363
QY	368 AACACCTGTTCAATTTGGTGTTCTTAAACAAGATTAATCTTCAAGCAAAATCACCTT 427
DB	364 AACACCTGTTCAATTTGGTGTTCTTAAACAAGATTAATCTTCAAGCAAAATCACCTT 423
QY	428 TTCAGSTGCATGCGATCTTCATTTCTCGTAAATATTTTGGATTCCTTAGAAGCATCTAG 487
DB	424 TTCAGSTGCATGCGATCTTCATTTCACTAATGATTAAGATTTTGGATTCCTTAGAAGCATCTAG 483
QY	488 TAAAAATTTCAAGCTACAAAGA TAGAAGGCTTTGGATTCGTTGATAGTGTTCAAAATTT 547
DB	484 TAAAAATTTCAAGCTACAAAGA TAGAAGGCTTTGGATTCGTTGATAGTGTTCAAAATTT 543
QY	548 AGTTGTTGAGAGAGAGGAACTATCATATGSCATATGCAAGATATGTTGGCCCAATGTTCTG 607
DB	544 AGTTGTTGAGAGAGAGGAACTATCAATATGSCAATGAGCAAGATATGTTGGCCCAATGTTCTG 603
QY	608 CAAAAATAATTAATATCATGCGCATGCGAGGATGCAACCAAGGCTTAACTTTTGGAAATTTG 667
DB	604 CAAAAATAATTAATATCATGCGCATGCGAGGATGCAACCAAGGCTTAACTTTTGGAAATTTG 663
QY	668 CAAAAATTTGANAAGTGAATATCTAAGAAGTAAAAAATGCAACAACAAATTCATATCAAAAT 727
DB	664 CAAAAATTTGANAAGTGAATATCTAAGAAGTAAAAAATGCAACAACAAATTCATATCAAAAT 723
QY	728 TGAATCATGCACTATGTTGTAGCTTCAAAATTTGATGATCAATGCTTCAGCAAAAGAGCCC 787
DB	724 TGAATCATGCACTATGTTGTAGCTTCAAAATTTGATGATCAATGCTTCAGCAAAAGAGCCC 783
QY	788 AAAATATCGATGGAATGATGATATCAATATCTCATATATTCCAAAATATCTGATACATATAT 847
DB	784 AAAATATCGATGGAATGATGATATCAATATCTCATATATTCCAAAATATCTGATACATATAT 843
QY	848 TGGAAACAGTATGATATTTCAATTTGTTTCGATCTCAAAATSTGAGGCGACAA 907
DB	844 TGGAAACAGTATGATATTTCAATTTGTTTCGATCTCAAAATSTGAGGCGACAA 903
QY	908 TATTACTTTGGTTCACAGCTCATGATATAGATTTGGAAGCTTAGATCTGGAATATTCAGA 967

Db 904 TATTACTTGGTCCAGGTCATGTTATTAAGTATTTGGAGCTTAGAGATCTGGAAATTCAGA 963
 QY 968 AGCTTATGTGTCTAATGTATCTGTAAATGAAGCCAAATATTCGGTCCGAAAAATGAGAT 1027
 Db 964 AGCTTATGTGTCTAATGTATCTGTAAATGAAGCCAAATATTCGGTCCGAAAAATGAGAT 1023
 QY 1028 TAGGATCAAGACTTGGCAGGAGAGATCTGCAACAGCTAGCAACATCTGAATTTCTGAATGT 1087
 Db 1024 TAGGATCAAGACTTGGCAGGAGAGATCTGCAACAGCTAGCAACATCTGAATTTCTGAATGT 1083
 QY 1088 GGAATGCAAGAGCTTAGTATCCCATATATATAGACCAAAATCTGTGATCGAGTTGA 1147
 Db 1084 GGAATGCAAGAGCTTAGTATCCCATATATATAGACCAAAATCTGTGATCGAGTTGA 1143
 QY 1148 ACCATGTATACAAAGTTTTCAGCAGTTCAAGTGAAGAAAAATGTGTATGAAATATCAA 1207
 Db 1144 ACCATGTATACAAAGTTTTCAGCAGTTCAAGTGAAGAAAAATGTGTATGAAATATCAA 1203
 QY 1208 GGGCACAAGTGCACAAAGGTGGCCATTAATTTGATTCGACGACAACTTTCATGTGA 1267
 Db 1204 GGGCACAAGTGCACAAAGGTGGCCATTAATTTGATTCGACGACAACTTTCATGTGA 1263
 QY 1268 AGGAATTAATATGAGATATTAATTTAGTAGGGGAAAGTGGAAAAACCATCAGAGGCTAC 1327
 Db 1264 AGGAATTAATATGAGATATTAATTTAGTAGGGGAAAGTGGAAAAACCATCAGAGGCTAC 1323
 QY 1328 GTGCAAAAATGTCAATTTTAAACATGCTGAACATGTTACACCACTGCATCTCACTGGA 1387
 Db 1324 GTGCAAAAATGTCAATTTTAAACATGCTGAACATGTTACACCACTGCATCTCACTGGA 1383
 QY 1388 AATTTCAAGAGATGAAGCTCTTTTGTATTAATTTATTAATTTATCTAGATCTTCAATAT 1447
 Db 1384 AATTTCAAGAGATGAAGCTCTTTTGTATTAATTTATTAATTTATCTAGATCTTCAATAT 1443
 QY 1448 ATAGCAGATATGATATATCACAATTAACAAATCTATATCTAGATTAATTAATTAAT 1507
 Db 1444 ATAGCAGATATGATATATCACAATTAACAAATCTATATCTAGATTAATTAATTAAT 1503
 QY 1508 TAAATATGACGATGAAGTTTAAATTAAGACTACTATGATTTCTATTTCTAGTCAAAA 1567
 Db 1504 TAAATATGACGATGAAGTTTAAATTAAGACTACTATGATTTCTATTTCTAGTCAAAA 1563
 QY 1568 GTTTGACGATGTACTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1625
 Db 1564 GTTTGACGATGTACTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1621
 RESULT 5
 AAS20853
 ID AAS20853 standard; DNA; 5822 BP.
 AC AAS20853;
 XX
 DT 05-JUN-2002 (first entry)
 DE Expression cassette, pFP-IRN1 useful for gene silencing.
 XX
 XX Inhibition of gene expression; expression cassette; inverted repeat; IR;
 XX NOS; nopaline synthase; figwort mosaic virus promoter; FMV; hsp70; PG;
 XX plant heat shock 70; tomato; polygalacturonase; disease resistance;
 XX flavour; nutritional characteristic; plant; gene silencing; pFP-IRN1;
 XX regulating gene expression; mutant; ds.
 XX
 XX Lycopersicon esculentum.
 OS Agrobacterium tumefaciens.
 OS Figwort mosaic virus.
 OS Synthetic.
 OS Chimeric.
 XX
 XX W0200214472-A2.
 PN
 XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US025538.
 PF 15-AUG-2000; 2000US-0225508P.
 ER 07-AUG-2001; 2001US-00924197.
 ER
 XX
 PA (DNAP) DNA PLANT TECHNOLOGY CORP.
 XX
 PI Guterson N, Oeller P;
 XX WPI; 2002-257599/30.
 DR
 XX
 XX
 PT Reducing the expression of a target gene in a cell, comprises expressing
 PT in the cell an expression cassette comprising a promoter operably linked
 PT to a sense or antisense targeting sequence and an inverted repeat of a
 PT nopaline synthase gene.
 XX
 XX
 PS Claim 51; Page 37-39; 39pp; English.
 CC
 CC The present invention relates to an improved method for inhibiting the
 CC expression of a target gene in a cell, by expressing in the cell an
 CC expression cassette comprising a promoter operably linked to a sense or
 CC antisense targeting sequence having substantial identity to a subsequence
 CC of the target gene, and an inverted repeat (IR) of a subsequence of a NOS
 CC (nopaline synthase) gene, where the IR is unrelated to the targeting
 CC sequence. The expression cassette, pFP-IRN1 is constructed using a
 CC figwort mosaic virus (FMV) promoter in which the 5'-untranslated leader
 CC (5'-UTR) is derived from a plant heat shock 70 (hsp70) gene, the tomato
 CC polygalacturonase (PG) gene and an IR of the terminator of the
 CC Agrobacterium tumefaciens NOS gene. The method is useful for regulating
 CC expression of endogenous genes and transgenes, e.g. to regulate
 CC expression of endogenous plant phenotypes such as disease resistance,
 CC flavour, protein or nutritional characteristics. The improved gene
 CC silencing construct is used in functional genomics to determine the
 CC effect of regulating gene expression of a selected endogenous gene or
 CC transgene. The method is simple and rapid, and is suitable for high-
 CC throughput studies. Multiple transgenic constructs all containing the
 CC same repeat element can be silenced at the same time, since the initial
 CC silencing trigger mediated through the inverted repeat region will apply
 CC to all of the transcripts. The present sequence represents the expression
 CC cassette, pFP-IRN1
 CC
 XX
 SQ Sequence 5822 BP; 1689 A; 1275 C; 1292 G; 1566 T; 0 U; 0 Other;
 Query Match 63.7%; Score 1042.4; DB 6; Length 5822;
 Best Local Similarity 99.4%; Pred. No. 2,1e-173;
 Matches 1046; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 376 TTCAATTTGTGTTCTTAAACAAAGATTAATCTTCTCAGCAAAATCACTTTTCAGGTC 435
 Db 1379 TTAACCAATGGTCTTAAACAAAGATTAATCTTCTCAGCAAAATCACTTTTCAGGTC 1438
 QY 436 CATGCAAGATCTCTAATTTCAATGAATTTTGAATCTCTTAAGCATCTAGTAAATTT 495
 Db 1439 CATGCAAGATCTCTAATTTCAATGAATTTTGAATCTCTTAAGCATCTAGTAAATTT 1498
 QY 496 CAGACTACAAAGATGAAGAGCTTTGATTTGATTAAGTGTCAAAATTTAGTTGTTG 555
 Db 1499 CAGACTACAAAGATGAAGAGCTTTGATTTGATTAAGTGTCAAAATTTAGTTGTTG 1558
 QY 556 GAGGAGAGAGAACTATCAATGTCATGCAACAGTAAGTGGCCAAATTTCTGGCAAAATTA 615
 Db 1559 GAGGAGAGAGAACTATCAATGTCATGCAACAGTAAGTGGCCAAATTTCTGGCAAAATTA 1618
 QY 616 ATAAATCACTGCATGCAAGGATGACCAACAGGCTTAACCTTCTGAATTTGCAAAATTT 675
 Db 1619 ATAAATCACTGCATGCAAGGATGACCAACAGGCTTAACCTTCTGAATTTGCAAAATTT 1678
 QY 676 TGAAGTGAATATATCAAGATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 735
 Db 1679 TGAAGTGAATATATCAAGATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1738
 QY 736 GCACATATGTTGAGCTTCAAAATTTGATGATCAATGCTTCAAGCAAGGCCAAATATCTG 795

PR	22-JUL-1999	99US-01550899
PR	22-JUL-1999	99US-0145152P
PR	23-JUL-1999	99US-0145145P
PR	23-JUL-1999	99US-0145221P
PR	26-JUL-1999	99US-0145218P
PR	27-JUL-1999	99US-0145276P
PR	27-JUL-1999	99US-0145913P
PR	27-JUL-1999	99US-0145918P
PR	28-JUL-1999	99US-0145919P
PR	02-AUG-1999	99US-0145951P
PR	02-AUG-1999	99US-0146386P
PR	02-AUG-1999	99US-0146388P
PR	04-AUG-1999	99US-0147038P
PR	04-AUG-1999	99US-0147204P
PR	04-AUG-1999	99US-0147192P
PR	05-AUG-1999	99US-0147130P
PR	05-AUG-1999	99US-0147260P
PR	06-AUG-1999	99US-0147303P
PR	09-AUG-1999	99US-0147416P
PR	09-AUG-1999	99US-0147493P
PR	10-AUG-1999	99US-0147935P
PR	11-AUG-1999	99US-0148171P
PR	12-AUG-1999	99US-0148319P
PR	13-AUG-1999	99US-0148341P
PR	13-AUG-1999	99US-0148365P
PR	16-AUG-1999	99US-0148684P
PR	17-AUG-1999	99US-0149175P
PR	18-AUG-1999	99US-0149168P
PR	20-AUG-1999	99US-0149426P
PR	20-AUG-1999	99US-0149712P
PR	20-AUG-1999	99US-0149723P
PR	23-AUG-1999	99US-0149922P
PR	23-AUG-1999	99US-0149930P
PR	23-AUG-1999	99US-0150566P
PR	26-AUG-1999	99US-0150884P
PR	27-AUG-1999	99US-0151065P
PR	27-AUG-1999	99US-0151067P
PR	27-AUG-1999	99US-0151080P
PR	30-AUG-1999	99US-0151303P
PR	31-AUG-1999	99US-0151338P
PR	01-SEP-1999	99US-0151930P
PR	07-SEP-1999	99US-0152633P
PR	10-SEP-1999	99US-0153070P
PR	13-SEP-1999	99US-0153786P
PR	15-SEP-1999	99US-0154018P
PR	16-SEP-1999	99US-0154039P
PR	20-SEP-1999	99US-0154779P
PR	22-SEP-1999	99US-0155139P
PR	23-SEP-1999	99US-0155486P
PR	24-SEP-1999	99US-0156563P
PR	28-SEP-1999	99US-0156458P
PR	29-SEP-1999	99US-0157117P
PR	04-OCT-1999	99US-0157753P
PR	05-OCT-1999	99US-0157865P
PR	06-OCT-1999	99US-0158026P
PR	08-OCT-1999	99US-0158232P
PR	12-OCT-1999	99US-0158239P
PR	13-OCT-1999	99US-0158293P
PR	13-OCT-1999	99US-0158294P
PR	14-OCT-1999	99US-0158295P
PR	14-OCT-1999	99US-0158330P
PR	14-OCT-1999	99US-0158331P
PR	14-OCT-1999	99US-0158638P
PR	18-OCT-1999	99US-0158584P
PR	21-OCT-1999	99US-0160741P
PR	21-OCT-1999	99US-0160761P
PR	21-OCT-1999	99US-0160766P
PR	21-OCT-1999	99US-0160770P

PR	21-OCT-1999;	99US-0160814E
PR	21-OCT-1999;	99US-0160815P
PR	22-OCT-1999;	99US-0160980P
PR	22-OCT-1999;	99US-0160981P
PR	22-OCT-1999;	99US-0160982P
PR	25-OCT-1999;	99US-0161404P
PR	25-OCT-1999;	99US-0161405P
PR	25-OCT-1999;	99US-0161406P
PR	25-OCT-1999;	99US-0161359P
PR	26-OCT-1999;	99US-0161360P
PR	26-OCT-1999;	99US-0161361P
PR	26-OCT-1999;	99US-0161970P
PR	28-OCT-1999;	99US-0161982P
PR	28-OCT-1999;	99US-0161993P
PR	29-OCT-1999;	99US-0162142P

Query Match	27.2%	Score	445.2	DB	3	Length	1393
Best Local Similarity	64.7%	Pred. No.	6.6e-69				
Matches	720	Conservative	0	Mismatches	368	Indels	24
						Gaps	3

QY	277	TTATGACTTACTGCTTTGGAGCTTAGGGTGATGGAAAAACATATGATATATTTGCATTTG	336
Db	300	TTAGTGTTCCTAACTTCGGAGCCAAAGAGATGGAAAACTGATGATACAGGCTTCA	359
QY	337	AGCAACATGGAATGAAGCATGTTCATCTTGAAACACCTGTCAATTGTGGTTCCTAAA	396
Db	360	AGAAACGATGGAAGAAAGCATGTTCACAAATGAGATTACTTTCTTGGTTCCTAAAG	419
QY	397	ACAAGAATTATCTTCAAGCAATCACCTTTTCAGGTCCATGCAGATCTTCAATTTCAG	456
Db	420	GAAGACTTATCTCTTAAGTCTACTCGATTGAGGGCCATCGAAATCTTACGTACT	479
QY	457	TAAAGTTTTGGATCCTTAGAAGCATCTGTGTAATTTTCAACTACAAAGATGAAGGC	516
Db	480	TTCAAGTCTTAGGCATTTATCAGCATCTACGAAAGCTTCAGATTACAAAGACAAAACC	539
QY	517	TTTGCATTCCTTTGATAGTGTCAAAATTTAGTTGTGGAGAGAGA-----GGACTA	570
Db	540	ATTGGCTTATCTTAGAGACGTTTAAACATCTATCAATCGACGGTGCCTGACGGGAATTA	599
QY	571	TCATATGCATGACACAAGATAGTGGCCAAAGTCTTGCAAAATTTAAATCACTGCAT	630
Db	600	TTAATGGCAACGAAAAAACTGGTGGCAGAACTCAATGCAAAATCGAANAATCAAGCAT	659
QY	631	GGAGGGATGACCAAGGGCTTAACCTCTGGAATTCGAAAAATTTGAAATGGAATATC	690
Db	660	GCACAAAAGCTCCAAAGGCTCTACTTTATTAACATTTTAAAGATTTGAATGTGMAAATC	719
QY	691	TAAAGATGTAATAATGCACACAAATTCATCATAAATTTGAGTCATGCACTAATGTTG	750
Db	720	TGAGGATGAAAAATGCGACGAGATTCAGATTTCAATTGAGAAATGCACAAAGTTGAAG	779
QY	751	CTTCAATTTGATGATCAATGCTTCAGCAAGAGCCCAATTCGTATGAGTGCATGAT	810
Db	780	TTAGTAATGTTAGATCACTGCTCCGGGCGATAGTCCCAACAGATGTATCCATATCA	839
QY	811	CAATATCTCAATATTTCAAAATATGATACATATATTGGAACAGGTGATGATTTGATTT	870
Db	840	CTAATACTCAAAACATTCGAGTCTCCAACTCGAATATCGGAACAGGTGATGATTAAT	899
QY	871	CAATGTTTCTGAGTCCAAAATGTGCAGGCCACAAATATCTTGTGGTCCAGGTCAG	930
Db	900	CCATTGAGAGATGGAAGCGAAAAATCTTCAAAATCTTGATTTAACTTGGCGGCCGCTCAG	959
QY	931	GTAATAGATTTGGAAGCTTTAGGATCTGAAATTCGAAAGCTTATGTGTCTAAATGTTACTG	990
Db	960	GGATAGAGATTTGGGAGCTTTGGGGAGCGCAAAATTCGAAAGCTTATGCTCGGGAATTAATG	1019
QY	991	TAAATGAAGCCAAAATTAATCGGTGCCGAAAATGAGCTTATAGATCAAGACTTGGCAGGGAG	1050
Db	1020	TGATAGTGTCTAAATCTCTGAGAGTGCACAAATGAGATTAGAGATTAAAGCTTATTCAGGGAG	1079
QY	1051	GATCTGACAAAGCTAGCAACATCAAAATTTCTGAATGTGGAATGCAACAGTTAATGATC	1110

QY 988 CTGTAATGAAGCCAAATTTATGCTGCCGAAATGGAGTTAGATCAAGACTTGGCAGG 1047
DB 1020 ATGTGATGTGTCTAGCGCTCTCTGAGACTGACATGAGTAAGATCAAGACTTACCAGG 1079
QY 1048 GAGGATCTGAGCAAGCTAGCAACATCAATTTCTGATGTGAAATGCAAGCGTTAAGT 1107
DB 1080 GAGGCTCAGGAACCTGCTAGAACTTTAAATTCGAAAATTCGATGATGATCAAGA 1139
QY 1108 ATCCATTAATTAAGACCAAACTATTGTGATGAGTTGACCATGTATCAACAGTTT 1167
DB 1140 ATCCGATCATATGACACAGAACTACTGCCA---CAAGGACAAATGCGAACAGCAGAAAT 1196
QY 1168 CAGCAGTTCAAGTAAATAATGTGTATAGAAATACAGGCGACAAGTSCAACAAAG 1227
DB 1197 CTGGGTTCAAGTAAACATGTCTGTATCAGAACTAAAGGTTACAGCGCAACAGATG 1256
QY 1228 TGGCCATTAATTTGATTTGACGACAAACTTTCCATGTGAGGAATTAATGAGAAATA 1287
DB 1257 TGGGATATATGTTTAATTTGAGTGTGAATATCCATGCAAGGTATGTGCTTGAGAAATG 1316
QY 1288 TAAATTTAGTAGGGGAAAGTGGAAAACCATCAGAGGCTACGTGCAAAAATGTCCATTTTA 1347
DB 1317 TGAACATCAAAAGGAGAA-----AAGCTTTTGGGAAAATGTCAATGTTA 1361
QY 1348 ACAATGCTGAACATGTTACACACACACTGCACTGACTAGAAATTTCA 1394
DB 1362 AGGATTAAGGCACTGTTCTCTTAATGCTTAATTAAGCTGA 1408

RESULT 10
AAZ22980
ID AAZ22980 standard; DNA; 1657 BP.
XX AAZ22980;
AC
XX 10-JAN-2000 (first entry)
DT
XX Brassica napus Sac66 protein encoding DNA.
DE
XX
KW Signal transduction protein; dehiscence; male sterile plant; D22 gene;
shatter resistance; oilseed rape; Sac66 protein; ss.
XX
OS Brassica napus.
XX
PN WO9949046-A1.
XX
PD 30-SEP-1999.
XX
PF 22-MAR-1999; 99WO-GB000905.
XX
PR 20-MAR-1998; 98GB-00006113.
XX
PA (BIOG-) BIOGENMA UK LTD.
PI Wyatt P, Roberts JA, Whitelaw C;
XX
DR WPI; 1999-580449/49.
XX
DR P-PSDB; AAY42649.
XX
PT A nucleic acid encoding a signal transduction protein involved in plant
dehiscence, useful for producing shatter resistant male sterile plants.
XX
PS Example 6; Fig 15; 71pp; English.
XX
CC The invention provides a nucleic acid encoding a signal transduction
protein involved in the process of dehiscence. The nucleic acids and
proteins are useful for regulating or controlling dehiscence of a pod or
an anther in a plant, useful in the production of shatter resistant or
shatter-resistant plants such as oilseed rape (Brassica napus). The present
sequence represents the DNA encoding a B. napus Sac66 protein
XX
SQ Sequence 1657 BP; 541 A; 313 C; 352 G; 451 T; 0 U; 0 Other;

Query Match 25.6%; Score 418.8; DB 2; Length 1657;
Best Local Similarity 62.4%; Pred. No. 2,8e-64;
Matches 736; Conservative 0; Mismatches 417; Indels 27; Gaps 4;
QY 277 TTAATGACTTACTTGTGAGGCTTAAGGTATGAAAACATATGATATATGATTCATTG 336
DB 350 TTAGTGTGTAACCTTGGAGCAAAAGGTATGAAAACGATGATCTCAGGCTTCA 409
QY 337 AGCAAGCATGAATGAAAGCATGTTTCATCTAGAAACACCTGTTCATTTGTGTTCTTAAA 396
DB 410 AGAAGCATGGAAGAGGACATGTTTCAACAAATGAAAGATGACTACTTCTTGTATCTTAAAG 469
QY 397 ACAAGATTTATCTTCTCAGCAAAATCACTTTTCAAGTCCATGCAATCTTCTATTTTAC 456
DB 470 GGAAGACTTATCTTCTTAAAGCTATGATTAAGAGGCCATGCAATCAATTAAGTACT 529
QY 457 TAAAGATTTTGTGATCCTTGAAGACATCTAGTAATTTCAAGTAC---AAAGATAGA 513
DB 530 TCCAGATCTTAGGCACTTTATCAGCTTCTACAAAACATGGAATTAAGTATGACAGA 589
QY 514 GGCCTTGGATTGCTTTGATAGTGTTCAAAATTTA-----GTTGTGAGAGAGAGAA 567
DB 590 ACCACTGCTTATTTTGAAGACGTTAAATATCTATCATGATGAGGCGCTCGGCGGGA 649
QY 568 CTATCAATGGCAATGACCAAGTATGTGGCCAGTTCTTGCAAAATTAATATCACTGC 627
DB 650 TTGTGATGGCAAGGAAAATCTGTGTGGCAAACTCATGCAAAATCACAATCTTAAGC 709
QY 628 CATGAGGGAATGACCAAGCGCTTAACTTCTGGAATGCAAAAATTTGAAAGTA 687
DB 710 CATGACAAAAGCCCAAGCGCTTACTCTCTACACCTTAAACAATTTGAAATGGAAGA 769
QY 688 ATCTAAGATGAATAATGACACACAATTCATATCAATTTAGTCATGACTAAATGTTG 747
DB 770 ATCTGAGTGAAGAAATGACAGAGATTCAGATTTGATGAGAAATGCAACAGTGTG 829
QY 748 TAGCTTCAATTTGATGATCAATGCTTACGAAAGGCCCAATCTGATGAGATCGATG 807
DB 830 ATGTTAAGATGTTAAGATCACTGCTCCGGGATGATGCCAAGATGATTCATA 889
QY 808 TATCAAAATCTCATATATTTCAAAATGCTATCTATTTAGGACAGGTGATGTTGA 867
DB 890 TCGTTGCTACTAAAACATTCGAATCTCAATTCAGACATTTGGACAGGTGATGTTGA 949
QY 868 TTTCAATTTGTTCTGATCTCAAAATGTCAGGCCCAAAATTAATCTGTGTCCAGTTC 927
DB 950 TATCCATTTGAGATGATCGCAAAATGTTCAAAATCAATGATTTAACTTGGCCCGGTC 1009
QY 928 ATGTTATAGTATTTGAAAGCTTTAGATCTGAAATTCAGAGCTTATGTCTAATGTTA 987
DB 1010 ATGGCATGAGCATTTGGAAGCTTGGGGATGACAAATTCGAAAGCTTATGTATGGGAATTA 1069
QY 988 CTGTAATGAAGCCAAATTTATCGGTGCGGAAAATGGAATGGAATCAAGACTTGGCAGG 1047
DB 1070 ATGTGATGTGTCTAGCGCTCTCTGAGACTGACATGAGTAAGATCAAGACTTACCAGG 1129
QY 1048 GAGGATCTGAGCAAGCTAGCAACATCAATTTCTGATGTGAAATGCAAGCGTTAAGT 1107
DB 1130 GAGGCTCAGGAACCTGTAAGACATTTAAATTCGAAAATTCGATGATGATCAAGA 1189
QY 1108 ATCCATTAATTAAGACCAAACTATTGTGATGAGTTGACCATGTATCAACAGTTT 1167
DB 1190 ATCCGATCATATGACCAAGAACTACTGCGA---CAAGGACAAATGCGAACCAAGAAAT 1246
QY 1168 CAGCAGTTCAAGTAAATAATGTGTATAGAAATACAGGCGACAAGTSCAACAAAG 1227
DB 1247 CTGGGTTCAAGTAAACATGTCTGTATCAGAACTAAAGGTTACAGCGCAACAGATG 1306
QY 1228 TGGCCATTAATTTGATTTGACGACAAACTTTCCATGTGAGGAATTAATGAGAAATA 1287
DB 1307 TGGGATATATGTTTAATTTGAGTGTGAATATCCATGCGCAAGGTATGTGCTTGAGAAATG 1366

QY 1288 TAAATTTAGTAGGGGAAAGTGGAAAAACCATCAGAGGCTACGTCGCAAAAATGTCATTTTA 1347
 DB 1367 TGAACATTAAGAGAGAA-----AAGCTTCTGCAAAAATGTCATTTTA 1411
 QY 1348 ACAATGCTGAACATGTTACACACACTGCTCACTAGAAATTTTCAGAGATGAAGCTC 1407
 DB 1412 AGCATAAAGGACCGTTTCTCTTAATGCCCTTAATTTCTAAGTTGATTAATGTAATAC 1471
 QY 1408 TTTTGTATTAATTTAATTTAATTTACTATAGATCTTCAATAT 1447
 DB 1472 ATAAATACGTATTAATGTTGTTATGATGACATCTATAT 1511

RESULT 11

ADA68538
 ID ADA68538 standard; DNA; 1280 BP.

AC ADA68538;

DT 20-NOV-2003 (first entry)

DE Arabidopsis thaliana gene, SEQ ID 554.

KW Plant; bacterial infection; fungal infection; viral infection; ds.

OS Arabidopsis thaliana.

PN W02003000898-A1.

PD 03-JAN-2003.

PF 22-JUN-2001; 2001WO-IB001105.

PR 22-JUN-2001; 2001WO-IB001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,

PI Katagiri F, Qian S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G,

DR WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

PS Claim 6; SEQ ID NO 554; 899bp; English.

CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

SQ Sequence 1280 BP; 401 A; 257 C; 292 G; 330 T; 0 U; 0 Other;

Query Match 23.9%; Score 391.6; DB 7; Length 1280;

Best Local Similarity 63.2%; Pred. No. 1.6e-59;

Matches 667; Conservative 0; Mismatches 359; Indels 30; Gaps 3;

QY 277 TTAATGTAAGTCTTAGAGCTAGGATGAGGATGAGAAAAATATGATATATGCAATTTG 336
 DB 203 TTAGTCTTTCGATTTTGAAGCTAAAGAGATGAGAAAAACCATGACACGAGCGTTTCG 262
 QY 337 AGCAAGCATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 396
 DB 263 TGAATGCTGAGAAAGCATGTTCTTCAATGAGAGCTGTTATCTCTAGTCTCTAAAG 322

QY 397 ACAGAAATATCTTCTCAAGCAAAATCACTTTGAGGCGCATGAGATCTTCAATTTGAG 456
 DB 323 GGAATTAATCTTCTTCAAGCTTCTTCAATTTAATCTGTCATGCAATTTCTTCAAGCG 382
 QY 457 TAAAGATTTTGGATCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 516
 DB 383 TTCAATATTTGGAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 442
 QY 517 TTTGATGCTTTTGAATGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 570
 DB 443 AATGATTAATGTTGATGAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 502
 QY 571 TCAATGCAATGCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 630
 DB 503 TGAACGAAACCGGAAACCTGCTGAGCAAACTCATGCAAAACGGAACAA----- 551
 QY 631 GCAAGGATGCAACCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 690
 DB 552 -----AGCTAAGCTCTTACTTTTCAAACTGCAAAAGCTGATGTAAGATC 601
 QY 691 TAAAGATTAATGCAACCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 750
 DB 602 TGAAGGTGAGAAATGCAACAGATCCAGATTTGATGAAAAATGCTCCAAAGTTGAGG 661
 QY 751 CTTCAATTTGATGATCAATGCTTCAAGCAAAAGCCCAATTAATCAATGATGAGTCAATG 810
 DB 662 TCTCTAATGCTGATGATCAATGCTTCAAGCAAAAGCCCAATTAATCAATGATGAGTCA 721
 QY 811 CAATTAATCAATTAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 870
 DB 722 CTAAACCCCAAAACATTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 781
 QY 871 CAATTTGCTGATGATCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 930
 DB 782 CTATTAAGGTGATCAAAATGTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 841
 QY 931 GATTAAGTATGGAACCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 990
 DB 842 GTATCAATTTGGAACCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 901
 QY 991 TAAATGACCAAAATTTATCGTCCGAAATGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1050
 DB 902 TGATGCTGCTAAGCTTCCGATGACCAATGATGATGATGATGATGATGATGATGATGAT 961
 QY 1051 GATCTGACAAAGCTGACCAATTTATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 1110
 DB 962 GATCAGGAACCTGCTGCAATTTATTTTCAAAACATTCAGATGATTAATGTTAAGATC 1021
 QY 1111 CCATTAATTTAGCAAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170
 DB 1022 CCAATTAATTTCAAAAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1078
 QY 1171 CAGTTCAAGTAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1230
 DB 1079 CGGTCAAGTAAAGAGCGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1138
 QY 1231 CCAATTAATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290
 DB 1139 CAATTAAGCTTAACTGACAGCAAGAACTATCCATGCAAGAAATGCTGACAGATGAT 1198
 QY 1291 ATTAGTGGGGAAGTGAAGAAACATCAAGAGCTTA 1326
 DB 1199 ACATTAAGGAGGAAAGCAACTTGCACCAATGCTA 1234

RESULT 12

ADA70539
 ID ADA70539 standard; DNA; 1269 BP.

AC ADA70539;

DT 20-NOV-2003 (first entry)

PS Example 174; Fig 8A-C; 68pp; English.

This is the nucleotide sequence of a tomato polygalacturonase genomic DNA clone isolated from a genomic library by screening with polygalacturonase cDNA. The polygalacturonase gene promoter is active in at least the breaker through red fruit stage in tomato fruit, and can be used in a novel method of the invention for producing a plant with an altered phenotype. In this method, regulatory regions from genes expressed during a particular developmental stage or in a specific tissue are identified by cDNA screening. The regulatory regions are manipulated for use with foreign sequences for introduction into plant cells to provide transformed plants having a phenotypic property that can be modulated. The invention is exemplified with light, seed and fruit-specific promoters. Transformation without gall formation of cells which are not susceptible to galling, such as tomato fruit, is achieved.

Apobacterium hosts, especially cereals and legumes, is achieved.

(Updated on 25-MAR-2003 to correct PF field.)

Sequence 2207 BP; 870 A; 282 C; 270 G; 784 T; 0 U; 1 Other;

Query Match	20.3%;	Score 331.8;	DB 2;	Length 2207;
Best Local Similarity	95.3%;	Pred. No. 4.7e-49;		
Matches 342;	Conservative 0;	Mismatches 17;	Indels 0;	Gaps 0;

Qy	1	AAATCTTTTCAATGACAAAGTTTAAAAACCATCCATATACATATATCATGTATCC	60
Db	1430	AAATCTTTTCAATGACAAAGTTTAAAAACCATCCATATACATATATCATGTATCC	1489
Qy	61	AAAGGAATAGATATCTCCTTCTCATATATATTTTGGCTTCATCAATTCAACTTGAGAA	120
Db	1490	AAAGGAATAGATATCTCCTTCTCATATATATTTTGGCTTCATCAATTCAACTTGAGAA	1549
Qy	121	GCAATGTTATATGATGACAAATTATTCAAACAAGTTATGATATATATTTCTTGACACAGAT	180
Db	1550	GCAATGTTATATGATGACAAATTATTCAAACAAGTTATGATATATATTTCTTGACACAGAT	1609
Qy	181	TTGCTCATGATTTTCAGCTTATCTTCTTATTTAGACAAAAATTTGAAAGCAACATA	240
Db	1610	TTGCTCATGATTTTCAGCTTATCTTCTTATTTAGACAAAAATTTGAAAGCAACATA	1669
Qy	241	ATATTGACAAAGTTGATATAAATGGGATTAAGTATTAATGTACTTAGCTTTGGAGCTA	300
Db	1670	ATATTGACAAAGTTGATATAAATGGGATTAAGTATTAATGTACTTAGCTTTGGAGCTA	1729
Qy	301	AGGGTGATGAGAAAAACATATGATATATTTGCAATTTGACCAAGCATGAGATGATCT	359
Db	1730	AGGGTGATGAGAAAAACATATGATATATTTGCAATTTGACCAATTTGGAATATATTTCT	1788

RESULT 15

ID AAV36973 standard; DNA; 2207 BP

AC	AAV636973;	
XX		
XX	25-MAR-2003	(revised)
DT	26-OCT-1998	(first entry)
XX		
DE	Polygalacturonase genomic DNA.	
XX		
XX	Transgenic plant; tissue-specific gene expression; promoter	
KW	fruit-specific expression; tomato; polygalacturonase; ss.	
XX		
OS	Lycopersicon esculentum.	
XX		
PN	US5753475-A.	
PD	19-MAY-1998.	
XX		
PF	10-AUG-1993;	93US-00105852.
XX		
PR	17-JAN-1985;	85US-00692605.
PR	31-JUL-1986;	86US-00893152.
PR	26-MAY-1987;	87US-000534363.

PR 28-JUN-1987; 87US-00078781.
PR 25-JAN-1988; 88US-00147751.
PR 15-MAR-1988; 88US-00168130.
PR 29-APR-1988; 88US-00188361.
PR 22-NOV-1988; 88US-00267665.
PR 21-MAY-1990; 90US-00526123.
PR 09-JUL-1990; 90US-00550804.
PR 14-SEP-1990; 90US-00582241.
PR 08-AUG-1991; 91US-00742834.

PA (CALJ) CALGENE INC.

PI Houck CM,

DR WPI; 1998-311403/27.

Transformation of plants - with regulatory sequence containing constructs for tissue specific expression of genes.

PS Example 17; Fig 8A-C; 67pp; English

This is the nucleotide sequence of a tomato polygalacturonase genomic DNA clone isolated from a genomic library by screening with polygalacturonase cDNA. The polygalacturonase gene promoter is active in at least the CC breaker through red fruit stage in tomato fruit, and can be used in a novel method of the invention for producing a plant with a regulatable phenotype. In this method, regulatory regions from plant genes (and T-DNA and Ti or Ri plasmids) are manipulated for use with foreign sequences for introduction into plant cells to provide transformed plants having a phenotypic property that can be modulated. The invention also exemplified with light, seed and fruit-specific promoters. Also claimed are methods for altering the phenotype of a plant to impact a desired CC tissue, modifying the genotype of a plant to impact a desired CC characteristic to fruit, modifying transcription in fruit tissue, and expressing a heterologous DNA sequence of interest specifically in fruit tissue. The method provides transformation without gall formation in CC plants which have historically not been *Agrobacterium* hosts. (updated on 25-MAR-2003 to correct PR field.)

Sequence 2207 BP; 870 A; 282 C; 270 G; 784 T; 0 U; 1 Other;

Query Match	20.3%	Score 331.8	DB 2	Length 2207
Best Local Similarity	95.3%	Pred. No. 4.7e-49		
Matches 342	0	Mismatches 17	Indels 0	Gaps 0

QY 1 AATCTTTTCATGACAAAGTTTAAAAACCTACCATATACATATATCATGTGTATCC 60

Db 1430 AATCTTTTCATGACAAAGTTTAAAAACCTACCATATATCATGTGTATCC 1489

QY 61 AAAGGAATAGATATCTCCTTCATATATTTTGGCTTCATCAATTTCAACTTGTA 120

Db 1490 AAAGGAATAGATATCTCCTTCATATATTTTGGCTTCATCAATTTCAACTTGTA 1549

QY 121 GCATGTATTGATGACAAATTTATTCAAACAAGTTATGATTAATTTCTTGACAAGAT 180

Db 1550 GCATGTATTGATGACAAATTTATTCAAACAAGTTATGATTAATTTCTTGACAAGAT 1609

QY 181 TTGCTCATGATTTTCAAGCTTATCTTCTTATTGAGCAAAAATTTGAAAGCAACATA 240

Db 1610 TTGCTCATGATTTTCAAGCTTATCTTCTTATTGAGCAAAAATTTGAAAGCAACATA 1669

QY 241 AATTTGACAAAGTTGATAAAAATGSGATTAAGATTAATGTAAGTTAGCTTTGGAGCTA 300

Db 1670 AATTTGACAAAGTTGATAAAAATGSGATTAAGATTAATGTAAGTTAGCTTTGGAGCTA 1722

QY 301 AGGGTGATGAAAAACATATGATTAATATTGCAATTTGACCAAGCATGATGAAGCATGT 359

Db 1730 AGGGTGATGAAAAACATATGATTAATATTGTAAGTATTTAAATATGGAATATATTGT 1788

Search completed: July 26, 2004, 23:42:35
Job time : 674 secs

QY 361 CATCTAGACACCTGTTCAATTGTGTTCCCTAAACACAGAATTATCTTCTCAGCAA 420

Db 361 CACTAGAACACCTGTCATTTGTGGTCTTAAACCAAGAAATATCTTCTCAAGCAAA 420
QY 421 TCACCTTTTCAAGGTCATGAGATCTTCTATTTCACTAAAGATTTTGGATTCCTTGAAG 480
Db 421 TCACCTTTTCAAGGTCATGAGATCTTCTATTTCACTAAAGATTTTGGATTCCTTGAAG 480
QY 481 CACTAGTAAATTTTCAGCTACCAAGATGAAAGCTTTGGATTTGCTTTGATAGTTC 540
Db 481 CACTAGTAAATTTTCAGCTACCAAGATGAAAGCTTTGGATTTGCTTTGATAGTTC 540
QY 541 AAAATTTAGTTGGTGGAGAGAGAACTATTCATGSCATGAGCAATGATGCTGCCAA 600
Db 541 AAAATTTAGTTGGTGGAGAGAGAACTATTCATGSCATGAGCAATGATGCTGCCAA 600
QY 601 GTTCTTGCAAAATTAATTAATCACTGCGCATGACAGGATGACCAACGGCCTTAACTTCT 660
Db 601 GTTCTTGCAAAATTAATTAATCACTGCGCATGACAGGATGACCAACGGCCTTAACTTCT 660
QY 661 GGAATTCGCAAAATTTGAAAGTGAATACTTAAAGATTAATGCAACCAATTCATA 720
Db 661 GGAATTCGCAAAATTTGAAAGTGAATACTTAAAGATTAATGCAACCAATTCATA 720
QY 721 TCAAAATTTGAGTCACTGACCTAAATGTTGAGCTCAAAATTTGATGATGATGCTCAGCA 780
Db 721 TCAAAATTTGAGTCACTGACCTAAATGTTGAGCTCAAAATTTGATGATGATGCTCAGCA 780
QY 781 AGAGCCCAAAATTAATGAGTCACTGATGATCAATTAATCAATTAATCAATTAATCTGATA 840
Db 781 AGAGCCCAAAATTAATGAGTCACTGATGATCAATTAATCAATTAATCAATTAATCTGATA 840
QY 841 CTATTTATGGAACAGGTGATGATTTGATTTCAATTTGCTGAGTCTCAAAATGTGACG 900
Db 841 CTATTTATGGAACAGGTGATGATTTGATTTCAATTTGCTGAGTCTCAAAATGTGACG 900
QY 901 CCACAAATTAATTAATGCTGAGTCACTGATGATTAAGTATGGAAGCTTGAAGTCTGGA 960
Db 901 CCACAAATTAATTAATGCTGAGTCACTGATGATTAAGTATGGAAGCTTGAAGTCTGGA 960
QY 961 ATTCAAGAGCTTATGCTCTATATGTTACTGTAATGAAAGCCAAATTAATGCTGCTGAAA 1020
Db 961 ATTCAAGAGCTTATGCTCTATATGTTACTGTAATGAAAGCCAAATTAATGCTGCTGAAA 1020
QY 1021 ATGAGATTAAGTCAAGCTTGGCAGGAGATCTGACAGCTTGAACCAATTCATCAATTC 1080
Db 1021 ATGAGATTAAGTCAAGCTTGGCAGGAGATCTGACAGCTTGAACCAATTCATCAATTC 1080
QY 1081 TGAATGTGGAATGCAAGCGTTAATGATCCATTAATTAATGACCAAAATCTATGTCATC 1140
Db 1081 TGAATGTGGAATGCAAGCGTTAATGATCCATTAATTAATGACCAAAATCTATGTCATC 1140
QY 1141 GAGTTGAACCATGTAACACAGTTTTCAGCAGTTCAAGTGAAGAAATGTGTGTATGAGA 1200
Db 1141 GAGTTGAACCATGTAACACAGTTTTCAGCAGTTCAAGTGAAGAAATGTGTGTATGAGA 1200
QY 1201 ATATCAAGGGCACAAGTGCACAAAGGTGGCCATTAATTTGATTCAGCACAACCTTTC 1260
Db 1201 ATATCAAGGGCACAAGTGCACAAAGGTGGCCATTAATTTGATTCAGCACAACCTTTC 1260
QY 1261 CATGTGAAGGAATTTAATGAGAAATATTAATTTGTGAGGGAAGTGGAAAACATCAG 1320
Db 1261 CATGTGAAGGAATTTAATGAGAAATATTAATTTGTGAGGGAAGTGGAAAACATCAG 1320
QY 1321 AGGCTACGTCGCAAAATGTCATTTTAAACATGCTGAACATGTTACACCACTGCACTT 1380
Db 1321 AGGCTACGTCGCAAAATGTCATTTTAAACATGCTGAACATGTTACACCACTGCACTT 1380
QY 1381 CACTAGAAATTTTCAAGAGTGAAGCTTTTGTATATTAATTAATTAATTAATTAATTAAT 1440
Db 1381 CACTAGAAATTTTCAAGAGTGAAGCTTTTGTATATTAATTAATTAATTAATTAATTAAT 1440
QY 1441 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
Db 1441 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500

Db 1441 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1500
QY 1501 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
Db 1501 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
QY 1561 GTCAAAAGTTTGAAGATGATGATTTTAAATGTAACAAATTAATTAATTAATTAATTAAT 1620
Db 1561 GTCAAAAGTTTGAAGATGATGATTTTAAATGTAACAAATTAATTAATTAATTAATTAAT 1620
QY 1621 GAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1636
Db 1621 GAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1636

RESULT 2
US-08-941-532-5
Sequence 5, Application US/08941532
Patent No. 6096946
GENERAL INFORMATION:
APPLICANT: ROBERTS, Jeremy Alan
APPLICANT: COUPE, Simon Allan
APPLICANT: JENKINS, Elizabeth Sarah
TITLE OF INVENTION: CONTROL OF POD DEHISCENCE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941,532
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00757
FILING DATE: 29-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9506684.1
FILING DATE: 31-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0623.0580001/RWE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1657 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 145..1446
US-08-941-532-5

Query Match 25.7%; Score 420.4; DB 3; Length 1657;
Best Local Similarity 62.5%; Pred. No. 2.8e-87;
Matches 737; Conservative 0; Mismatches 416; Indels 27; Gaps 4;

QY 277 TTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 336
Db 350 TTAATGTTTGAACCTTGGAGAGCAAAAGGTGATGGAAGAAACCGATGATCTCAGGCTTTCA 409

QY 337 AGCAAGATGAGTGAAGATGCTACCTAGAACACCTGTTCAATTTGGCTTCTTAAA 396
 DB 410 AGAAGATGGAAGAGGATGTCACAAATGAGTGAAGTCTTCTTGTATCTCTTAAAG 469
 QY 397 ACAGAAATTATCTCTCAAGCAAAATCACCTTTGAGGCCATGAGATCTTCTATTTGAG 456
 DB 470 GGAAGACTTATCTCTTAAAGCTATTTAGATTTAGAGGCCATGCAAAATCATTAAGTACT 529
 QY 457 TAAAGATTTTGGATCTTGAAGCAATCTAGTAAATTTTGAAGTAC---AAAGATGAA 513
 DB 530 TCCAGATCTAGGCACTTTATCAGCTTCTAATAAATCATGGATTAAGTAATGACAGA 589
 QY 514 GGGTTTGGATTCCTTTGATAGTGTCAAAATTTA-----GTTGTGAGAGAGAGAA 567
 DB 590 ACCACTGGCTTATTTTGGAGAGCTTAAATCATCATGATGAGGGCTCGCGGGGGA 649
 QY 568 CTATCAATGGCAATGAGCAAGTATGAGGCCAAGTTTCCAAAATTAATTAATCACTGC 627
 DB 650 TTGTTGATGGCAACGAAAAATCTGGTGGCAAACTCATCAAAATGCAAAATCTTAAGC 709
 QY 628 CATGCAAGGATGCAACACGGCTTAACTTCTGGAATGCAAAATTTGAAAGTGAATA 687
 DB 710 CATGCAAAAAGGCCAAGGCTCTTAATCTTAACAACCTAAACATTTGAATGTAAGA 769
 QY 688 ATCTAAAGATTAATGCAACAATTTATATCAATTTGATGATGCACTAATGTTG 747
 DB 770 ATCTGAGATGAAATGACACAGATTCAGATTTGATGAAATGCAACAGTGTG 829
 QY 748 TAGCTTCAATTTGATGATCAATGCTTCAGCAAAAGGCCAATTAATGATGAGTGCATG 807
 DB 830 ATCTTAAGATGTTAAGATCTGCTCTGCGCATGATCCCAACAGGATGATTCATA 889
 QY 808 TATCAAAATCTCAATATTAATTAATCTGATCTATTAATTTGAAACAGTATGTTGA 867
 DB 890 TCGTTCCTACTAAAACATTCGATCTCAATCAACATTTGGAACAGTATGATGCA 949
 QY 868 TTTTCAATTTTCTGATCTCAAAATGTCAGGCCCAAAATTAATCTTGTGTCCAGGTC 927
 DB 950 TATCCATTAAGATGATGCAAAATGTTCAATCATGATTTAACTTGGCCCGCGTGC 1009
 QY 928 ATGTAATAGTATTTGGAAGCTTGAAGTCTGGAATTCAGAAAGTTATGCTCTAAAGTTA 987
 DB 1010 ATGGATCTGCAATTTGGAAGCTTGGGAGATGACATTTCAAGCTTAATGATGGAATTA 1069
 QY 988 CTGTAATGAAGCCAAATTAATCGTGCAGAAATGAGTTGATCAAGACTTGGCAGG 1047
 DB 1070 ATGTGATGCTGCTACGCTCTCTGAGACTGACAAATGAGATTAAGATCAAGACTTACAGG 1129
 QY 1048 GAGGATCTGGACAAAGCTAGCAACATCAATTTCTGATGTAAGTGAAGTCAAGCGTTAAGT 1107
 DB 1130 GAGGATCTGGAAGCTGCTAAGAACATTAATCAAAACATGCTGATGATATGTCAGA 1189
 QY 1108 ATCCCATATTTAGAACCAAACTATTTGATGAGTTGAACCAATGTAACAAGCTTTT 1167
 DB 1190 ATCCCATATTAATCAACGAACATCTGCGA---CAAGCAAAATGCGAACAACAGAT 1246
 QY 1168 CAGCATCTCAAGTAAAAATGCTGTATGAGATATCAAGGCAACATGCAACAAAGG 1227
 DB 1247 CTGCGATCTCAAGTAAACATGCTGTATGCGAACATCAAGTACGAGCGCAAGGATG 1306
 QY 1228 TGGGCATTAATTTGATTTGACACAACTTTCTCATGTAAGAAATTAATATGAGAAATA 1287
 DB 1307 TGGGCATTAATTTGATTTGACAGTGTGAATATCAATGCAAGTATTTGCTTGAATG 1366
 QY 1288 TAAATTTAGTGGGAAAGTGAACCAATCAAGAGCTAAGTGAACAAATGCTCTTTTA 1347
 DB 1367 TGAACATCAAGAGAGAA-----AAGCTTCTTGAACAAATGCTCAATGTTA 1411
 QY 1348 ACAATGCTGAACATGTTACACCACTGCACTTCACTGAATAATTCAGAGATGAAAGTCT 1407
 DB 1412 AGGATAAAGGACGCTTCTCTTAATGCTCTTAATTTCTAAGTTGATTAATATATAC 1471

QY 1408 TTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1447
 DB 1472 ATAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1511

RESULT 3
 US-09-051-239A-1
 / Sequence 1, Application US/09051239A
 / Patent No. 6420628
 / GENERAL INFORMATION:
 / APPLICANT: ULVSKOV, Peter
 / APPLICANT: CHILD, Robin
 / APPLICANT: VAN ONCKELIN, Henri
 / APPLICANT: PRINSEN, Els
 / APPLICANT: BORKHARDT, Bernard
 / APPLICANT: SANDER, Lilli
 / APPLICANT: PETERSEN, Morten
 / APPLICANT: BONDGARD POUlsen, Gert
 / APPLICANT: BOTTERMAN, Johan
 / TITLE OF INVENTION: Seed Shattering
 / FILE REFERENCE: 2121-0138P
 / CURRENT APPLICATION NUMBER: US/09/051,239A
 / PRIOR FILING DATE: 1998-09-28
 / PRIOR APPLICATION NUMBER: PCT/EP96/04313
 / PRIOR FILING DATE: 1996-10-04
 / PRIOR APPLICATION NUMBER: EP 95 402241.4
 / PRIOR FILING DATE: 1995-10-06
 / PRIOR APPLICATION NUMBER: EP 95 203328.0
 / PRIOR FILING DATE: 1995-12-08
 / NUMBER OF SEQ ID NOS: 14
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 1
 / LENGTH: 1631
 / TYPE: DNA
 / ORGANISM: Brassica napus
 / FEATURE:
 / OTHER INFORMATION: Location 95-163 = region encoding the presumed
 / OTHER INFORMATION: endo-PG signal peptide.
 / OTHER INFORMATION: Location 884-908 = region of the endo-PG cDNA
 / OTHER INFORMATION: corresponding to oligonucleotide PG3
 / OTHER INFORMATION: Location 1059-1073 = region of the endo-PG cDNA
 / OTHER INFORMATION: complementary to oligonucleotide PG2
 / OTHER INFORMATION: Location 1229-1245 = region of the endo-PG cDNA
 / OTHER INFORMATION: complementary to oligonucleotide PG5
 / NAME/KEY: CDS
 / LOCATION: (95)..(1393)
 / OTHER INFORMATION: Location 821-837 = region of endo-PG cDNA
 / OTHER INFORMATION: corresponding to oligonucleotide PG1.
 / OTHER INFORMATION: Strain cv. Topaz.
 / NAME/KEY: unsure
 / LOCATION: (1439)
 / OTHER INFORMATION: n = a, c, g, t, any, other, unknown, or other
 / US-09-051-239A-1

Query Match 25.7%; Score 420.2; DB 4; Length 1631;
 Best Local Similarity 63.6%; Pred. No. 3,1e-87;
 Matches 717; Conservative 0; Mismatches 383; Indels 27; Gaps 4;

QY 277 TTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 336
 DB 300 TTAGTGTTCGAACCTTCGAGCCAAAGAGATGAAAAACGATGATCTAGGCTTCA 359
 QY 337 AGCAAGATGGAATGAACATGTTGATCTAGAACCTGTTCAATTTGGTGGTTCTTAAA 396
 DB 360 AGAAGCATGGAAGAGGCAATGTTCAACAAATGAGTTACTACTTTTAATTCCTAAAG 419
 QY 397 ACAGAAATTATCTCTCAAGCAAAATCACCTTTTCAAGTCAATGAGATCTTATTTGAG 456
 DB 420 GAAAGACTTATCTCTTAAAGCTATTTAGATTTAGAGGCCATGCAAACTTTAAGTACT 479
 QY 457 TAAAGATTTTGGATCTTGAAGCAATCTAGTAAATTTTGAAGTAC---AAAGATGAA 513
 DB 480 TCCAGATCTAGGCACTTTATCAGCTTCTACAAAACGATCGAATTAAGTATGACAGA 539

```

QY 514 GGCCTTGAGTCTTTGATAGTTCATAATTTA-----GTTGTGAGAGAGAGAGAA 567
DB 540 ACCACTGGCTTATTTGGAGAGCATTAATATCATCAATGATGAGCGGCTCGGCGGGA 599
QY 568 CTATCAATGGCAATGACAGATATGAGTGGCAAGTTCTTGCAAAATTAATTAATCACTGC 627
DB 600 TTGTTGATGGCAACGAAATATCTGTGGGCAAAATCATGCAAAATGACAAATCTTAAG 659
QY 628 CATGCAAGGATGACCAAGCGCTTAACTTTGGAATTGCAAAATTTGAAAGTGATA 687
DB 660 CATGCAAAAAGCGCCCTCTTACTCTCAACAATAAGAAATTTGATGTGAGA 719
QY 688 ATCTAAAGATGAAAAATGCAACAATTCATATCAAAATTTGAGTCATGACCTAATGTTG 747
DB 720 ATCTGAAGATGAGAAATGACAGAGATTCAGATTTGATGAGAAATGCAACAATGTTG 779
QY 748 TAGCTTCAAAATTTGATGATCAATGCTTCAGCAAAAGCGCCAAATATCATGATGAGTCCATG 807
DB 780 GCGTTAAGATGTTAAGATCATCTCTCTGCGCATATGCCCCAACAGGATGTTATTCATA 839
QY 808 TATCAAAATCAATATATTCATTAATTCATATCTATTTTGGAAACAGTATATGTTGA 867
DB 840 TCGTTGCTACTAAATAATTCGAATTCGAATTCAGACATTTGAGAGAGTATGATGTTA 899
QY 868 TTTCAATTTGTTCTGCAATCTCAAAATGTCAGGCGCAAAATATTAATTTGTCAGGTC 927
DB 900 TATCATTTGAGATGATCGCAAAATGTTCAAAATCAATGATTTAACTTTGCGGCCCCGCTC 959
QY 928 ATGTATATGATTTGGAAGCTTAGATCTGGAATTCGAAGCTTATGTGTATGTTA 987
DB 960 ATGGATCAGATTTGGAAGCTTTGGGCGAGTGAACAATTCGAAGCTTATGTATGGAATTTG 1019
QY 988 CTGTAATGAAAGCCAAATTAATCGGTGCGGCAAAATGAGTTAGATCAAGCTTGGCAGG 1047
DB 1020 ATGTGATGATGCTACGCTCTCTGAGATGACAAATGAGTAAGATCAAGACTTAACAGG 1079
QY 1048 GAGGATCTGCAACAGCTGCAACATCAATTTCTGAATGTGAAATGCAAGCGTTAAGT 1107
DB 1080 GAGGCTCGAGAACTGCTAAGAACATTAATTTCCAAAATTCGTTGATGATATGTCAGA 1139
QY 1108 ATCCATATATATAGACCAAACTATTTGATGATGAGTGAACCATGATATACAAAGTTT 1167
DB 1140 ATCCGATATATATGACCAAGACTATCTGGA---CAAGGACAAATGCGCAACGAAAGAT 1196
QY 1168 CAGCAATTCAGTGAAGAAATGTTGTATGAGATATCAAGGCGCAAGTGCAGAAAG 1227
DB 1197 CTGCGGTTCAAGTGAACATGCTGTATGACAAATTAAGGTACGAGCGCAACAGATG 1256
QY 1228 TGGCCATTAATTTGATTTGAGCAACAACCTTCCATGTAAGGATTTAAATGAGAGATA 1287
DB 1257 TGGCATATATGTTAATTTGAGTGTGAAATATTCATGCGCAAGTATTTGCTTGAGATG 1316
QY 1288 TAAATTTAGTAGGGAAGTGAAGAAACATCAGAGGCTACGTGCAAAAATGTCATTTTA 1347
DB 1317 TGAACATCAAAAGAGAA-----AAGCTTCTTGCAAAAATGTCATATGTTA 1361
QY 1348 ACAATGCTGAACATGTTACACACACTGCACTTCTACTAGAAATTTCA 1394
DB 1362 AGGATTAAGGCACTGTTCTCTTAATGCGCTTAATTAAGCTGA 1408

```

```

RESULT 4
US-08-467-023-139
; Sequence 139, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griffith, Irwin J.;
; APPLICANT: Pollock, Joanne;
; APPLICANT: Bond, Julian F.;
; APPLICANT: Garman, Richard D;
; APPLICANT: Kuo, Mei-Chang;
; APPLICANT: Yeung, Siu-mei H.;

```

```

; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; TITLE OF INVENTION: Japanese Cedar Pollen
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Maltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
;
; Query Match 13.9%; Score 228; DB 3; Length 1410;
; Best Local Similarity 52.3%; Pred. No. 3.4e-43;
; Matches 557; Conservative 0; Mismatches 495; Indels 12; Gaps 2;

```

```

QY 240 AATATTGACAAGTGTGATTAATAATGAGTAAATGTAAGCTTTGAGCT 299
DB 4 AAAGTTGAGCAATCTCGTCATGATCTATCAACATCTTCAATGTGAAAGATGCGCA 63
QY 300 AAGGCTGATGAAAAACATATGATTAATTTGATTTGAGCAAGCATGCAATGAT 359
DB 64 GTAGGCGATGGAAGAACATGATGCACTGAGGATTTTCAACAGCATGGCAAGCTGATGC 123
QY 360 TCATCTAAGAACACCTGTTCAATTTGTTGTTCTTAATAAACAGATTAATCTCAAGCA 419
DB 124 AAAAAACATCAAGCAATGTTGCTTG---CGAGCAGCAAGAAATTTGTTTAACAT 180
QY 420 ATCACTTTTCAAGTGCATGACATCTTCTATTTTCAAGTAAGATTTTGGATCTTAGAA 479
DB 181 CTGTTCTTCAATGAGGCAATGTCACCTCACTTATTTAAGTATGATGGAATATAGCT 240
QY 480 GCATCTAGTAATAATTTCAAGCTTCAAAAGATGAGAGCTTTGATGCTTTGATGTT 539
DB 241 GCGTACCAAAATCCAGCGAGCTGGAAGATATAGAAATATGTTTCAAGTTCTAAACTT 300
QY 540 CAAATTTTATGTTTGGAGGAGAGAACTATCATGCGCAATGCACTATGATGAGGCA 599
DB 301 ACAGGTTTACTCTAATGGAATGGAAGGTATTAATTTGAGCAAGAAACAAATGTTGGCT 360
QY 600 AGTTCTTGCAAAATTAATTAATCACTGCCA-----TGAGGAGATCACCACAGGCC 650
DB 361 GGCATATGTAATGGTCAATGAGAGAAATTTGCAAGATGTGATATGACCAAGGCC 420

```

QY 651 TTAACTCTGGAATGGAAAATTTGAAAGTGAATATCTTAAGAGTAAATGCA 710
| | | | |
Db 421 ATTAATTCATTTTTCACAGGGCTGATATCCAGAGACTGAATGAAGCTGCC 480
| | | | |
QY 711 CAAATTCATATCAATTTGAGTCATGCACTTAATGTTAGCTCAAAATTTGATATCAT 770
| | | | |
Db 481 GAATTTCAATTTAGTTTGGAAATTTGAGAGAGTAAATATCATGGCATTTAGATTACG 540
| | | | |
QY 771 GCTTCAGCAAGAGCCCAAAATCTGATGAGTCCATGATCAAAATCTCATATATTTCAA 830
| | | | |
Db 541 GCACCGAGACAGCTCTTAACACATGATGAAATGATATCTTTGATCTAAATCTTAC 600
| | | | |
QY 831 ATATTCGATCTATTTATGAGACAGGATGATTTGATTTTCAATTTGATCTCA 890
| | | | |
Db 601 TTTCAAAAGACACGATAGAGACAGGGAGTACCTGCTCTATAGCACAGGGCTCTTCT 660
| | | | |
QY 891 AATGTCAGGCGCAAAATATTACTTGTGCTCCAGGTATGATGATTAAGTATGAGAGCTTA 950
| | | | |
Db 661 AATATTGTGATTTAGAGATCTGATTTGGGCTCCAGGCTATGAAATAGTATAGAGATCTT 720
| | | | |
QY 951 GATCTGGAATTCAGAGCTTATGCTCTAATGTTACTGTAATGAGCCAAATTTATC 1010
| | | | |
Db 721 GGGAGGAAAACCTCTAGACAGAGGTTTCATACGTGACGTAATGAGGCTAAATTCATA 780
| | | | |
QY 1011 GGTCCGAAAATGAGTTGAGATCAAGACTTGGCAGAGGAGATCTGACACAGCTAGAAC 1070
| | | | |
Db 781 GACACCAAAATGATTTAGAAATCAAAACATGGCAGGGTGGTTACAGCATGGCAACCAT 840
| | | | |
QY 1071 ATCAAAATTTCTGATGATGGAATGCAAGAGCTTAAGTATCCATTAATTTAGCCAAAC 1130
| | | | |
Db 841 ATATATTATGAGATATGTTGAATATGATTAATTCGAGAACCCCATATTAATTAATCAATTC 900
| | | | |
QY 1131 TATTTGATCGAGTTAACCAGTATPACAAAGTTTTCAGCAGTTCAAGTAAATTTG 1190
| | | | |
Db 901 TACTGACATCGACTTCTGCTGCAAAAACAGAGGCTCCGGTTCAATTCAGAAATG 960
| | | | |
QY 1191 GTGATGAGATATCAAGGCAAGTCAAGTCAACAAAGTGGCCATAAATTTGATTCAGC 1250
| | | | |
Db 961 AATACCAAGACATACGTGGGACATCAGCAACGACAGCAATTCACCTTAAGTCACT 1020
| | | | |
QY 1251 ACAAACTTTCATGTAAGAAATTTATATGAGAAATATAATTT 1294
| | | | |
Db 1021 GACAGTATGCCCTGCAAAATATTAAGCTAAGTATATCTTT 1064
| | | | |

RESULT 5
US-08-467-023-141
; Sequence 141, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griftech, Irwin J.;
; APPLICANT: Pollock, Joanne;
; APPLICANT: Bond, Julian F.;
; APPLICANT: Garman, Richard D;
; APPLICANT: Kuo, Mei-Chang;
; APPLICANT: Yeung, Siu-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; TITLE OF INVENTION: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025,6 USD2 (IMI-028CPD2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1479 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-467-023-141

Query Match 13.9%; Score 228; DB 3; Length 1479;
Best Local Similarity 52.3%; Pred. No. 3,4e-43;
Matches 557; Conservative 0; Mismatches 495; Indels 12; Gaps 2;

QY 240 AATATTGACAAAGTTGATATAAAATGGAATTAAGTATTAATGATTAATGAGCT 299
| | | | |
Db 73 AAAGTGGACATTTCTGTCATGATGATCTATCAACATCTTCAATGTGAAAAGTATGGCGCA 132
| | | | |
QY 300 AAGGATGAGAAAACATATGATATATTTGATTTGAGACATGAAATGAGCATGT 359
| | | | |
Db 133 GTAGGAGATGAAAGCATATGTCATGAGCAATTTTCAACACACATGGAAGCTGATGC 192
| | | | |
QY 360 TCATCTAGAACACCTGTTCAATTTGTTGCTCTTAAACAGAAATTAATCTTCAAGCA 419
| | | | |
Db 193 AAAAACCATCGACATATGTTGCTTG---CAAGGACAGAAATTTGTTGAACAAT 249
| | | | |
QY 420 ATCACTTTTCAGTTCATGACAGATCTTATTTTCAGTAAAGATTTTGGATCTTAGAA 479
| | | | |
Db 250 CTGTTCTTCATGAGGCGCATGTCACCTCACTTTAATTTAAGTATGAGGATTAATAGCT 309
| | | | |
QY 480 GCATCTAGTAAATTTCACTACAAAGTAAAGTCTTATTTGATTAATGTT 539
| | | | |
Db 310 GCGTACCAAAATTCACGAGCTGGAAGATTAATGATGGCCAGGAAAACATATGTTGGCT 369
| | | | |
QY 540 CAAAATTTAGTTTGGAGAGAGAGAACTATCAATGCGCATGCAATGATGAGGCGCA 599
| | | | |
Db 370 ACAGTTTACTCTAATGGGTAAAGTGTAAATTTGATGGCCAGGAAAACATATGTTGGCT 429
| | | | |
QY 600 AGTTCTTGCAAAATTAATTAATCACTGCCA-----TGAGGAGTACACCAAGGCC 650
| | | | |
Db 430 GGCATATGTAATGAGTCAATGATGAGAGAAATTTGCAACGATGATGATGACCAACAGCC 489
| | | | |
QY 651 TTAACCTCTGGAATTTGCAAAAATTTGAAAGGATTAATCTAAAGATTAATGACAA 710
| | | | |
Db 490 ATTAATTCGATTTTTCACAGGGCTGATATCCAGGACGTAAGCTAAATGAAACAGTCCC 549
| | | | |
QY 711 CAAATTCATATCAATTTGAGTCATGCACTAATGTTGATCTTCAAAATTTGATGATCAAT 770
| | | | |
Db 550 GAATTTCAATTTAGTTTGGGATTTGAGAGAGTAAATATCATGGCATTTAGTATTAAG 609
| | | | |
QY 771 GCTTACGAAAGAGCCCAAAATCTGATGAGTCAATGATCAAAATCTCAATATATTTCAA 830
| | | | |
Db 610 GCACGAGACAGCTCTTAACACATGATGAAATGATATCTTTGATCTTAATAAATCTTAC 669
| | | | |
QY 831 ATATCTGATCTATTTATGAGAAAGGATGATGATTTCAATGTTTCTGATCTCA 890
| | | | |
Db 670 TTACAAAAGACAGATGAGAACAGGGATATCTGCTGCTATAGGCAAGAGGCTCTCT 729
| | | | |
QY 891 AATGTCAGGCGCAAAATATTACTTGTGTCAGGTCATGATATGATTTGGAAGCTTA 950
| | | | |

Db	730	AAATTTGTGATTGAGGATCTGATTTGGGGTCCAGGCCATGGATTAAGTATAGGAAGTCTT	789
Qy	951	GGATCTGGAAATTCAGAGCTTATGTCCTAATGTTACTGTAAATGAAGCCAAATTTATC	1010
Db	790	GGAGGGAAAACTTAGAGCAGAGAGTTTCAATACGTGACGTAATTTGGGGCTAAATTCATA	849
Qy	1011	GGTCCGAAAAATGAGTTTAGGATCAAGCTTGCGCAGGGAGGATCTGCACAAGCTACGAC	1070
Db	850	GACACACAAATGAGATTAGAGATCAAAAACATGGCAGGGTGTTGACGAGATGGCAAGCAT	909
Qy	1071	ATCAAAATTTCTGAATGTGGAAATGCAAGACGTTTAAGTATCCCATTAATTATGACCAAAAC	1130
Db	910	ATAATTTATAGAAATGTGAAATATATAATTTGGAGAACCCCATTTAATTAATCAATTC	969
Qy	1131	TATTGTATGAGTTGAAACCATGATATACAAGTTTCAGAGTTCAAGTAAATAATGTG	1190
Db	970	TACGCACTTCAGGTTCTGCTTGCCAAAACCGAGAGTGTGGGTTCAANTCCAAAGATGTG	1029
Qy	1191	GTCATATGGAATATCAAGGGCAGACAGTGCACAAGAAGTGCGCATTAATTTGATTTGCAGC	1250
Db	1030	ACATTTCAAGAACATACGTGGAGCAATCGCAACACAGCAGCAATTCACACTTAAGTGCAGT	1089
Qy	1251	ACAACTTTCATGTGAAGAAATTAATGAGATATATAATTT	1294
Db	1090	GACGATGCCCTGCAAAGATATTAAGTAAAGTATATATCTTT	1133

RESULT 6
 US-08-467-023-133
 ; Sequence 133, Application US/08467023
 ; Patent No. 6090386
 GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.;
 APPLICANT: Pollock, Joanne;
 APPLICANT: Bond, Julian F.;
 APPLICANT: Garman, Richard D;
 APPLICANT: Kuo, Mei-Chang;
 APPLICANT: Yeung, Siu-wei H.;
 APPLICANT: Brauer, Andrew;
 APPLICANT: Exley, Mark A.;
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 261
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 STREET: 610 Lincoln St
 CITY: Waltham
 STATE: MA
 COUNTRY: USA
 ZIP: 02154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467, 023
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Remillard
 REGISTRATION NUMBER: 38,872
 REFERENCE/DOCKET NUMBER: 025.6 US2D (IMI-028CPD2)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 133:
 SEQUENCE CHARACTERISTICS:

```

: LENGTH: 1726 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: CDNA
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: 42..1586
:
:
US-08-467-023-133

```

Query Match	13.9%;	Score 228;	DB 3;	Length 1726;
Best Local Similarity	52.3%;	Pred. No. 3.6e-43;		
Matches 557;	Conservative 0;	Mismatches 495;	Indels 12;	Gaps 2

QY	240	AAATATGACAAAGGTGATAAAAAATGGATTAAGATTAATTAAGTACTGCTTGGAGCT	299
Db	180	AAAGTTGAGCACTTCGTCACTGATGCTATCCAAATCTCAAGTGGAAAAAGTATGGCGCA	239
QY	300	AAGGTGATGGAAAAACATATGATPAATATTGCACTTTGAGCAGCATGGAATGAAGATGT	359
Db	240	GTAGGCGCATGGAAAAAGCATGATGTCACCTGAGGCAATTTTCAACACAGATGGCAAGCTGATGC	299
QY	360	TCATCTGAGAACCTGTTGCAATTTGGTGTTCCPATAAAACAAGATTTATCTTCCAGCAA	419
Db	300	AAAAACCCATCAGCAATGTTGCTTGTC--CCAGGACAGCAAGAAATTTGTTGAACATAT	356
QY	420	ATCACCTTTTCAGGTCCATGACAGATCTTCTATTTACGTAAGATTTTGGATCTTGA	479
Db	357	CTGTCTTTCATGGGCGCATGTCAACCTCACTTACTTTAAGGTAGATGGATTAATAGCT	416
QY	480	GCATCTGTATAAATTTTACACTCAAGATAAGAAAGGCTTTGGATGCTTTTGTATGTTGTT	539
Db	417	GCGTACCAAAATCCAGGAGCTGGAGAAATATGATATGTTGTCAGTTTGCTAACTT	476
QY	540	CAAAATTTAGTGTGTGGAGAGAGAGCAACTATCATGTCGAATGSAACAATATGTTGGCCA	599
Db	477	ACAGGTTTACTCTAATAGGTTAAAGGTGTAATGATGGGCAAGAAACAATGTTGGGCT	536
QY	600	AGTTCCTTGCAAAATTAATTAATCACTGCCA-----TGCAGGATGCAACCAAGGCC	650
Db	537	GGCCAAATGATAATGGGTCAATGAGCGAGAAATTTGCACAGATGCTGATGACCAACAGCC	596
QY	651	TTAACCTTTCGMAATTCGCAAAAATTTTAAAGTGAATATCTAAAGATTAATAATATGACAA	710
Db	597	ATTAAATTCGATTTTTCACGGGCTGTATATCCAAAGATGGAACCTAATAGAACAGTCCC	656
QY	711	CAAAATCATATCAAAATTTGAGTCATGCACTPAATGTTTACCTTCAAAATTTGATGATCAAT	770
Db	657	GAAATTCATTTAGTTTGTGGAAATTTGAGGAGATAAAAATCAATCGGCATTAAGTATACG	716
QY	771	GCTTCAGCAAAAGCCCAATACGTAGTGAAGTGCATGATCAATACTCAATPATATTCAA	830
Db	717	GCACCGAGAGCAATCTTACACCTGATGAAATTTGATATCTTTGCATCTAATAAACTTTAC	776
QY	831	ATATCTGATACTAATTTATGGAACAGGTGATGATTTGATTTCAATTTGTTCTGCATCTCAA	890
Db	777	TTACAAAAGAACAGATAGGAACAGGGAGATGACGTGCTCCCTATAGGCAACAGGATCTTCT	836
QY	891	AATGTGCAGGCGCAAAATTTACTTGTGTGTCACAGTCAATGATGTAATTAATTTGGAAGCTTA	950
Db	837	AAATATTGATTGATGGAGATCTGATTTTGGGTCACAGGCAATGGAATATGATATAGAAAGTCTT	896
QY	951	GGATCTGGAATTTTCAAGAAGCTTATGTGTCTATGTTACTGTAATTAAGGCCAAATTTATC	1010
Db	897	GGGAGGAAAACTTAAGCAGAGAGTTTCAACGTGCACGTAATATGGGGCTAAATTCATA	956
QY	1011	GGTGCGGAAAAATGAGTTTAGATCAAGCTTTGGCAGGAGAGATTTGCAACAAGCTATGCAAC	1070
Db	957	GACACACAAAATGAGTTAAAGATCAAAAACATGCGAGGATGGTTGACGATGCGAAGCAT	1016
QY	1071	ATCAAAATTTTGAATGTGGAATGCAAGACGTTAAGTATCCCATPATTTATGACCAAAAC	1130
Db	1017	ATAATTATTAAGAAATGTTGAATGATTAATTTGGAGAAACCCCATATTTAATTAATCAATTC	1076

Patent No. 5759829
GENERAL INFORMATION:
APPLICANT: SHEWMAKER, C.
APPLICANT: KRIDL, J.
APPLICANT: HIATT, W.
APPLICANT: KNAUF, V.
TITLE OF INVENTION: ANTI-SENSE REGULATION OF GENE
EXPRESSION IN PLANT CELLS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,213
FILING DATE: 5-JUNE-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/750,505
FILING DATE: 27-AUG-91
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/240,408
FILING DATE: 30-AUG-88
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/920,574
FILING DATE: 17-OCT-86
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/845,676
FILING DATE: 28-MAR-86
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 26-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 219 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-463-213-1

Query Match 13.0%; Score 212; DB 1; Length 219;
Best Local Similarity 100.0%; Pred. No. 8.4e-40;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 ATGGATTAAAGTGAATTAATGCTTGAAGCTAAGGGTATGGAACATATG 321
DB 8 ATGGATTAAAGTGAATTAATGCTTGAAGCTAAGGGTATGGAACATATG 67

QY 322 ATAAATTGCAATTGAGCAAGCATGGAATGTTCAATGAACACCTGTTCAAT 381
DB 68 ATAAATTGCAATTGAGCAAGCATGGAATGTTCAATGAACACCTGTTCAAT 127

QY 382 TTGTGTTCTTAATAAACAAGATTATCTTCTAAGCAATCACTTTTCAAGTCCATGCA 441
DB 128 TTGTGTTCTTAATAAACAAGATTATCTTCTAAGCAATCACTTTTCAAGTCCATGCA 187

QY 442 GATCTTCTATTTCAGTAAGATTTTGGATCC 473
DB 188 GATCTTCTATTTCAGTAAGATTTTGGATCC 219

RESULT 9
5453566-1
Patent No. 5453566
APPLICANT: SHEWMAKER, CHRISTINE K.; KRIDL, JEAN C.; HIATT, WILLIAM R.; KNAUF, VIC
TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
IN PLANT CELLS
NUMBER OF SEQUENCES: 2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/750,505
FILING DATE: 27-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 240,408
FILING DATE: 30-AUG-1988
APPLICATION NUMBER: 920,574
FILING DATE: 17-OCT-1986
APPLICATION NUMBER: 845,676
FILING DATE: 28-MAR-1986
SEQ ID NO: 1
LENGTH: 219

Query Match 13.0%; Score 212; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 8.4e-40;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 ATGGATTAAAGTGAATTAATGCTTGAAGCTAAGGGTATGGAACATATG 321
DB 8 ATGGATTAAAGTGAATTAATGCTTGAAGCTAAGGGTATGGAACATATG 67

QY 322 ATAAATTGCAATTGAGCAAGCATGGAATGTTCAATGAACACCTGTTCAAT 381
DB 68 ATAAATTGCAATTGAGCAAGCATGGAATGTTCAATGAACACCTGTTCAAT 127

QY 382 TTGTGTTCTTAATAAACAAGATTATCTTCTAAGCAATCACTTTTCAAGTCCATGCA 441
DB 128 TTGTGTTCTTAATAAACAAGATTATCTTCTAAGCAATCACTTTTCAAGTCCATGCA 187

QY 442 GATCTTCTATTTCAGTAAGATTTTGGATCC 473
DB 188 GATCTTCTATTTCAGTAAGATTTTGGATCC 219

RESULT 10
US-08-941-532-7
Sequence 7, Application US/08941532
Patent No. 6096946
GENERAL INFORMATION:
APPLICANT: ROBERTS, Jeremy Alan
APPLICANT: COUPE, Simon Allan
APPLICANT: JENKINS, Elizabeth Sarah
TITLE OF INVENTION: CONTROL OF POD DENSIFICATION
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941,532
FILING DATE: 30-SEP-1997

Query Match	4.5%	Score 73;	DB 3;	Length 164;
Best Local Similarity	67.3%;	Pred. No. 5.6e-08;		
Matches 103;	Conservative 0;	Mismatches 50;	Indels 0;	Gaps 0;
QY	786	CCAAATAGTGTGAGGCCATGATTCAAATACTCAATATATTCATATCTGATCTATT	845	
Db	10	CCGATACGACGGATTCATATCTGTTGCTACTAAAACTTCGATCTCCAAATCGAC	69	
QY	846	ATTGGACAGGTGATGATTGTATTTCAATGTTCTTCGATCCAAAATGTGACGCCACA	905	
Db	70	ATTGGACAGGTGATGATTGCATATCCATTGAGATGATGCAAAATGTTCAATCAAT	129	
QY	906	AATATTACTGTTGGTCCAGTCGATCGATTAAGT	938	
Db	130	GATTTAACTTGCGGCCCGGTCACGGCCCTAAGT	162	

```

? FILING DATE: 26-AUG-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: BENT, Stephen A.
? REGISTRATION NUMBER: 29,768
? REFERENCE/DOCKET NUMBER: 30472/114 IMMU
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703)836-9300
? TELEFAX: (703)683-4109
? TELEX: 899149
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 7218 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? CLONE: pTZpc-F15
? IS-08-232-463-14

```

```

Query Match      3.6%; Score 58.4; DB 1; Length 7218;
Best Local Similarity 5.0%; Pred. No. 0.00045;
Matches 20; Conservative 22; Mismatches 158; Indels 0; Gaps 0;

QY    930 GGTAATAAGTATGGAACCTTAGATCTCGAAATTGCAGAAGCTTATGTCTAATGTTACT   989
       |||||
Db    1452 GATAGAAAGATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1395
                                     : :::: : :::: : :::: : :::: :

QY    990 GTAAATGAGACCCAAATTAATCGGTCCGAAAATGAGAGTTAGATCAACTGGCAGGA   1049
       : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: :
Db    1392 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1333
                                     : :::: : :::: : :::: : :::: :

QY    1050 GCATCTGCAGACAACCTGCAACATCAAAATTTCTGATATGGAATGCAAGCGTTAAGTAT   1109
       : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: :
Db    1332 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1273
                                     : :::: : :::: : :::: : :::: :

QY    1110 CCATATATTATAGACCAAACTATGTGATCGAGTTGAACATGATACAAACAAGTTTCA   1169
       : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: :
Db    1272 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1213
                                     : :::: : :::: : :::: : :::: :

QY    1170 GCAGTTCAAGTGAAAAATGTGTGTATGAAATATCAAGGCAACAOTGCAACAAGGTG   1229
       : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: :
Db    1212 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1153
                                     : :::: : :::: : :::: : :::: :

QY    1230 GCCATAAATTGATGTCAGACACAACCTTCATGTGAAGGAATTATATGAGAAATATA   1289
       : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: :
Db    1152 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1053
                                     : :::: : :::: : :::: : :::: :

QY    1290 AATTGTAGTGGGAAAGTGGAAAACATCAGAGCGTACGT   1329
       : :::: : :::: : :::: : :::: : :::: : :::: : :::: : :::: :
Db    1092 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRATGCAAGCTCCT   1053
                                     : :::: : :::: : :::: : :::: :

RESULT 12
US-09-790-988-1
: Sequence 1, Application US/09790988
: Patent No. 6632935
: GENERAL INFORMATION:
: APPLICANT: SHIGENOBU, SHUITI
: APPLICANT: WATANABE, HIDEKI
: APPLICANT: HATTORI, MASAHIRA
: APPLICANT: SAKAKI, YOSHIYUKI
: TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
: FILE REFERENCE: 081356/0159
: CURRENT APPLICATION NUMBER: US/09/790,988
: CURRENT FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: JP2000-107160
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 1
: LENGTH: 640681
: TYPE: DNA
: ORGANISM: Buchnera sp.
```


Db 609553 ATCCAGATTAAATTTTATTTATGATTTTCATATTTTTCANPAATTAATAATTTT 609494
QY 1453 AGATATGATATATACGATTAACAAATCTATATCTATGATGAAATTAATTAATA 1512
Db 609493 TATTACCTTAACCTTGATTTTAATA-----AGATAAAAAGAAATTTTATTAATTA 609439
QY 1513 TGTAAGATTTGAAGTTTAAATAGACTAGATTTCTATTTCTATGTCGAAAAGTTG 1572
Db 609438 ATTAATGATTTTATTTTATTTATTTCTATATTAATATGTTAATTAATAACAAAAATTTT 609379
QY 1573 ACGATTTGACTTTTAAATGTAACAAATTAATAAATGTTAATTAATAAAAAAAA 1632
Db 609378 TTGATTAGAAATATATTATTTCTAGAAAAATTTAAATATTTATATACGATTAATAATA 609319
QY 1633 AAAA 1636
Db 609318 AATA 609315

RESULT 15

US-10-204-708-19/C
; Sequence 19, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 19
; LENGTH: 6866
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-19

Query Match 3.2%; Score 51.8; DB 4; Length 6866;
Best Local Similarity 49.8%; Pred. No. 0.014;
Matches 131; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 1370 ACACGCACTTCACAGAAATTTGAGAGATGAAGCTCTTTGTATATTAATTAATTAT 1429
Db 2406 ACACGCACTTCACAGAAATTTGAGAGATGAAGCTCTTTGTATATTAATTAATTAT 1429
QY 1430 ACTATGATCTTCATATATATAGACATATGATATATGACATTAACAAATCTATATCTAT 1489
Db 2346 AATAAATACTATATATTAATAATATATATTTATATATATATATATTAATACTATATAT 2287
QY 1490 GTATTGAATTAATTAATTAATATATAGATGAAAGTTTAAATAGACTATATGTAAT 1549
Db 2286 TTTCTAAATATTTCTCTAAATATCCAAAAATTAATAATTAATAATTTCTTATACA 2227
QY 1550 TCTATTTCTAGTGAAGTTTGAGAGATTTGACTTTTAAATGTAACAAAAATTAATAAATG 1609
Db 2226 TCTCTAATAATTAATAATTAATAACATAAATTAATAATAAATACTAATATATATA 2167
QY 1610 GTTATTTATATGAAAAAAA 1632

Db 2166 TTATTATATCTAAAAAACAA 2144

Search completed: July 27, 2004, 04:23:19
Job time : 135 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2004, 23:42:40 ; Search time 763 Seconds

(without alignments)
10481.334 Million cell updates/sec

Title: US-10-018-604-1

Perfect score: 1636

Sequence: 1 aaccttttcaatagacaaag.....atatgaaaaaaaaaaaaa 1636

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3216467 seqs, 244149694 residues

Total number of hits satisfying chosen parameters: 6432934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1042.4	63.7	5822	13	US-09-924-197-1
2	420.2	25.7	1631	14	US-10-151-668-1
3	375	22.9	1359	17	US-10-437-963-24608
4	369	22.6	1182	17	US-10-437-963-24607
5	331.8	20.3	2207	13	US-09-782-130-26
6	281.2	17.2	1168	16	US-10-260-238-618
7	278.2	17.0	1335	17	US-10-437-963-28561
8	267.4	16.3	1308	13	US-10-425-114-30261
9	230	14.1	1809	17	US-10-437-963-24606
10	210.8	12.9	1333	13	US-10-424-599-61964
11	198	12.1	1219	13	US-10-425-114-10970
12	194.8	11.9	1479	13	US-10-424-599-141807
13	175	10.7	442	9	US-09-770-444-874
14	175	10.7	442	9	US-09-924-035A-569

15	171.2	10.5	1673	13	US-10-362-091-3	Sequence 3, Appli
16	169	10.3	1678	13	US-10-424-599-116781	Sequence 16781, A
17	168.6	10.3	1382	13	US-10-425-114-29325	Sequence 29325, A
18	164	10.0	870	13	US-10-424-599-59408	Sequence 59408, A
19	163.6	10.0	1984	13	US-10-424-599-92601	Sequence 92601, A
20	162	9.9	1631	13	US-10-425-114-9445	Sequence 9445, A
21	161	9.8	811	16	US-10-260-238-3398	Sequence 3398, A
22	154	9.4	596	17	US-10-021-323-9854	Sequence 9854, A
23	153.2	9.4	1472	13	US-10-424-599-95670	Sequence 95670, A
24	151.8	9.3	1185	9	US-09-938-842A-2520	Sequence 2520, Ap
25	151.8	9.3	1185	11	US-09-938-842A-2520	Sequence 2520, Ap
26	150.4	9.2	1272	17	US-10-437-963-14307	Sequence 14307, A
27	149.8	9.2	1067	13	US-10-425-114-16916	Sequence 16916, A
28	148.6	9.1	1374	17	US-10-437-963-43683	Sequence 43683, A
29	141.8	8.7	2012	13	US-10-424-599-33624	Sequence 33624, A
30	140.2	8.6	1432	13	US-10-425-114-7729	Sequence 7729, Ap
31	130.6	8.0	589	17	US-10-021-323-5922	Sequence 5922, Ap
32	130.6	8.0	746	9	US-09-770-149-31	Sequence 31, Appl
33	129.2	7.9	1248	16	US-10-260-238-498	Sequence 498, App
34	128.2	7.8	568	17	US-10-021-323-12284	Sequence 12284, A
35	126.6	7.7	2179	17	US-10-437-963-56617	Sequence 56617, A
36	125.6	7.7	565	13	US-10-424-599-83751	Sequence 83751, A
37	123	7.5	588	17	US-10-021-323-9731	Sequence 9731, Ap
38	121.6	7.4	1605	17	US-10-437-963-50404	Sequence 50404, A
39	120.2	7.3	602	17	US-10-021-323-5860	Sequence 5860, Ap
40	120.2	7.3	602	17	US-10-021-323-9178	Sequence 9178, Ap
41	118	7.2	1886	17	US-10-437-963-31774	Sequence 31774, A
42	117.8	7.2	1723	17	US-10-437-963-73025	Sequence 73025, A
43	117.6	7.2	588	17	US-10-021-323-9998	Sequence 9998, Ap
44	115	7.0	588	17	US-10-021-323-5864	Sequence 5864, Ap
45	115	7.0	588	17	US-10-021-323-9185	Sequence 9185, Ap

ALIGNMENTS

RESULT 1
US-09-924-197-1
Sequence 1, Appl: Application US/09924197
Publication No. US20030018993A1
GENERAL INFORMATION:
APPLICANT: Guttersen, Neal
TITLE OF INVENTION: Improved Methods of Gene Silencing Using Inverted
TITLE OF INVENTION: Repeat Sequences
FILE REFERENCE: 012176-010810US
CURRENT APPLICATION NUMBER: US/09/924,197
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: 60/225,508
PRIOR FILING DATE: 2000-08-15
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 5822
TYPE: DNA
ORGANISM: Agrobacterium tumefaciens
US-09-924-197-1

Query Match	63.7%	Score 1042.4;	DB 13;	Length 5822;
Best Local Similarity	99.4%	Pred. No. 3.5e-192;		
Matches 1046;	Conservative	0;	Mismatches 6;	Indels 0; Gaps 0;
QY	376	TTCAATTTGGTCTCTAAACAGAGTATCTCTCAAGCAATACCTTTTCAGGTC	435	
DB	1379	TTAAACATGATCTCTAAACAGAGTATCTCTCAAGCAATACCTTTTCAGGTC	1438	
QY	436	CATGAGATCTCTCAATTTCAAGTAAATTTTGAATCTTGAACATCTAATAATT	495	
DB	1439	CATGAGATCTCTCAATTTCAAGTAAATTTTGAATCTTGAACATCTAATAATT	1498	
QY	496	CAGACTCAAAAGATGAAGGCTTTGATGCTTTGATGATGATCAAAATTAGTTGTTG	555	
DB	1499	CAGACTCAAAAGATGAAGGCTTTGATGCTTTGATGATGATCAAAATTAGTTGTTG	1558	

```

QY 556 GAGAGAGAGAACTATCATATGCAATGACACAGTATGCGCAAGTTCTTGCAAAATA 615
DB 1559 GAGAGAGAGAACTATCATATGCAATGACACAGTATGCGCAAGTTCTTGCAAAATA 1618
QY 616 ATAAATATCTGCGATGAGAGGATGACCAACGGCTTAACCTTGGAATTCGCAAAAT 675
DB 1619 ATAAATATCTGCGATGAGAGGATGACCAACGGCTTAACCTTGGAATTCGCAAAAT 1678
QY 676 TGAAGATGATTAATCTAAAGATTAATAATGACACAAATTCATATCAATTTGAGTCAT 735
DB 1679 TGAAGATGATTAATCTAAAGATTAATAATGACACAAATTCATATCAATTTGAGTCAT 1738
QY 736 GCACTAAATGTTGATGCTTCAAAATTTGATGATCAATGCTTGACAAAGGCCAAATCTG 795
DB 1739 GCACTAAATGTTGATGCTTCAAAATTTGATGATCAATGCTTGACAAAGGCCAAATCTG 1798
QY 796 ATGAGATGCTATGATCAAAATCTCAATATATTTCAAAATCTGATATTAATTTGAAACAG 855
DB 1799 ATGAGATGCTATGATCAAAATCTCAATATATTTCAAAATCTGATATTAATTTGAAACAG 1858
QY 856 GTGATGATGTTGATTTCAATTTGCTGATCTCAAAATGTCAGGCCACAAATATTAATT 915
DB 1859 GTGATGATGTTGATTTCAATTTGCTGATCTCAAAATGTCAGGCCACAAATATTAATT 1918
QY 916 GTGATGATGTTGATTTCAATTTGCTGATCTGGAAGCTTGAATTCGAAAGCTTATG 975
DB 1919 GTGATGATGTTGATTTCAATTTGCTGATCTGGAAGCTTGAATTCGAAAGCTTATG 1978
QY 976 TGTCTAATGTTGATTTCAATTTGCTGATCTGGAAGCTTGAATTCGAAAGCTTATG 1035
DB 1979 TGTCTAATGTTGATTTCAATTTGCTGATCTGGAAGCTTGAATTCGAAAGCTTATG 2038
QY 1036 AGACTTGGCAGGAGATCTGGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1095
DB 2039 AGACTTGGCAGGAGATCTGGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 2098
QY 1096 AAGAGCTTAAGTATCCATTAATTAATTAAGACCAAACTATTTGATGATGAAACATGTA 1155
DB 2099 AAGAGCTTAAGTATCCATTAATTAATTAAGACCAAACTATTTGATGATGAAACATGTA 2158
QY 1156 TACCAACATTTTTCAGCAGTTCAGTGAAGAAATGCTGATGAGAAATTCGAAAGCTTGA 1215
DB 2159 TACCAACATTTTTCAGCAGTTCAGTGAAGAAATGCTGATGAGAAATTCGAAAGCTTGA 2218
QY 1216 GTGCAACAAAGGTGCGCATTAATTAATTTGATGATGAGCAAACTTTCATGATGAAAGATTA 1275
DB 2219 GTGCAACAAAGGTGCGCATTAATTAATTTGATGATGAGCAAACTTTCATGATGAAAGATTA 2278
QY 1276 TAATGAGAAATTAATTAATTTGATGAGGAGAAATGGAAGAAACCATGAGGCTTACGTCAAAA 1335
DB 2279 TAATGAGAAATTAATTAATTTGATGAGGAGAAATGGAAGAAACCATGAGGCTTACGTCAAAA 2338
QY 1336 ATGTCATTTTAAACAATGCTGAACATGTAACACACACTGCACTTCACTAGAAATTTTCAG 1395
DB 2339 ATGTCATTTTAAACAATGCTGAACATGTAACACACACTGCACTTCACTAGAAATTTTCAG 2398
QY 1396 AGGATGAAGCTCTTTTGTATTAATTTTAATTT 1427
DB 2399 AGGATGAAGCTCTTTTGTATTAATTTTAATTT 2430

```

RESULT 2

```

US-10-151-668-1
; Sequence 1, Application US/10151668
; Publication No. US20020184660A1
; GENERAL INFORMATION:
; APPLICANT: ULVSKOV, Peter
; APPLICANT: CHILD, Robin
; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els
; APPLICANT: BORKHARDT, Bernard
; APPLICANT: SANDER, Lilli

```

```

; APPLICANT: PETERSEN, Morten
; APPLICANT: BUNDEGARD POULSEN, Gert
; APPLICANT: BOTTERMAN, Johan
; TITLE OF INVENTION: Seed Shattering
; FILE REFERENCE: 2121-0138P
; CURRENT APPLICATION NUMBER: US/10/151,668
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US/09/051,239
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: PCT/EP96/04313
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: EP 95 402241.4
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: EP 95 203328.0
; PRIOR FILING DATE: 1995-12-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1631
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; LOCATION: 95-163 = region encoding the presumed
; LOCATION: (95)..(1393)
; OTHER INFORMATION: endo-Pg signal peptide.
; OTHER INFORMATION: Location 884-900 = region of the endo-Pg cDNA
; OTHER INFORMATION: complementary to oligonucleotide PG5
; OTHER INFORMATION: corresponding to oligonucleotide PG3
; OTHER INFORMATION: Location 1059-1073 = region of the endo-Pg cDNA
; OTHER INFORMATION: complementary to oligonucleotide PG2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(1393)
; OTHER INFORMATION: Location 1229-1245 = region of the endo-Pg cDNA
; OTHER INFORMATION: complementary to oligonucleotide PG5
; OTHER INFORMATION: Location 821-837 = region of endo-Pg cDNA
; OTHER INFORMATION: corresponding to oligonucleotide PGI.
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1439)
; OTHER INFORMATION: Strain cv. Topaz.
; OTHER INFORMATION: n = a, c, g, t, any, other, unknown, or other
; US-10-151-668-1
; Query Match 25.7%; Score 420.2; DB 14; Length 1631;
; Best Local Similarity 63.6%; Pred. No. 1.4e-71;
; Matches 717; Conservative 0; Mismatches 383; Indels 27; Gaps 4;
QY 277 TTAATGTAAGTACTGCTTGGAGCTTAAGGTGATGGAAGAAACATATGATATATTCGATTTG 336
DB 300 TTAATGTAAGTACTGCTTGGAGCTTAAGGTGATGGAAGAAACATATGATATATTCGATTTG 359
QY 337 AGCAGCATGGAATGAGCATGTCATCTAGAACACCTGTCATATTTGCTTCTTAA 396
DB 360 AGAAGCATGGAAGGAGCATGTCATCTAGAACACCTGTCATATTTGCTTCTTAA 419
QY 397 ACNAGATTAATCTTCTCAAGCAATCACTTTTCAGGTTCATGCAATCTTCTATTTGAG 456
DB 420 GAAAGCTTAATCTTCTTAAAGTCTATTAATGATGAGAGGCCCATGCAATCTTAACTGAGCT 479
QY 457 TAAAGATTTTGGATTCCTTAGAAGCATCTAGTAAATTTTCAAGCTC---AAAGATAGAA 513
DB 480 TCAAGATCTTGAAGCATTTATGAGCTTTCACAAAGCATGAGATTCAGATTAATGACAAG 539
QY 514 GGGTTGATGCTTCTTGAATGATGTTCAAAATTTA-----GTTGTTGAGAGAGAGAA 567
DB 540 ACCACTGCTTATTTTGGAGAGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 599
QY 568 CTATCAATGCAATGAGCAAGTATGTTGGCAAGTTCTTGCAAAATTAATTAATTAATTA 627
DB 600 TTGTTGATGCAACGGAATATCTGTTGGCAAAATTCATGCAAAATTCGAAATTCGAAAG 659
QY 628 CATGCAAGGATGACCAACGAGCTTAACTTTCGAAATTCGAAATTTGAAAGTGAATA 687
DB 660 CATGCAACAAAGGCGCAACGAGCTTAACTTTCGAAATTCGAAATTTGAAAGTGAATA 719

```

```
QY 688 ATCTAAGAGTAAATTCACACAAATTCATCAATTTGAGTCATGCACTAATGTTG 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 720 ATCTGAGAGTGAAGAAATGACAGACAGATTGATTTGCATTTGAGAAATGCAACATGTTG 779
QY 748 TAGCTTCAAAATTTGATGATCAATGCTTCAGCAAGAGCCCAAAATATCTGATGAGATCCATG 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 GCGTTAAGAAATGTTAAGATCACTGCTCTGCGCATAGTCCCAACACCGGATGTAATTCATA 839
QY 808 TATCAAAATCTCAATATATATTCAAATATCTGATCACTATTAATTTGAAACAGTGATGTTGA 867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 840 TCGTTGCTACTAAAAACATTCGAATCTCCAAATTCAGACATTTGGACAGGATGATATTGTA 899
QY 868 TTTCAATTTGTTTCTGGAATTCCAAAATGTCAGGCCCAAAATTTACTTTGTTGTCAGATC 927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 900 TATCCATTTGAGAGATGATCCGCAAAATGTTCAAAATCAATGATTTAACTTTGCGGCCCGGTC 959
QY 928 ATGATTAATGATTTGGAAGCTTAGGATCTGGAATTCGAAATTCGAAAGCTTAGTGTCTATGTTA 987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 960 ATGGATCAGCATTTGGAAGCTTGGGGATGACAAATTCGAAAGCTTATGTAATCGGGAATTTG 1019
QY 988 CTGTAAATGAGCCCAAAATTTATCGTCCGCAAAATGAGTTAGATCAAGCTTGGCAGG 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1020 ATGTGATGTTGCTACGCTCTTGAGACTGCAATGAGTAAGATCAAGATTTACAGG 1079
QY 1048 GAGGATCTGACAGAGCTAGACACATCAATTTCTGAAATGGAATGCAAGAGCTTAAAT 1107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1080 GAGGGTCAAGAACTGCTAGAGAACATTTAAATTCGAAACATTCGTATGATATATGTCAGA 1139
QY 1108 ATCCCAATATTTATAGACCAAAACATATGATGAGTGAACCATGATATCAACAGTTT 1167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1140 ATCCGATCATATACGACCAAGACTACTGCGA---CAAGACAAATGGAACACACAGAT 1196
QY 1168 CAGCAGTTCAAGTGAATAATGTGATGATGAGAAATATCAAGGCACAAGTGCACAAAG 1227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1197 CTGCGGTTCAAGGAAACAAATGTGTATATGAAACATTAAGGTAAGAGGCAACAGATG 1256
QY 1228 TGGCCATAAAATTTGATTGACACAAACTTTCATGTGAAAGAAATTAATGAGAGATA 1287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1257 TGGCGTAATGTTTAATTTGAGTGTGAAATATCCATGCCAAAGTATTTGCTTGAGAGATG 1316
QY 1288 TAAATTTAGTAGGGGAAAGTGAAAAACATCAGAGCTGCGTGCAAAAAATGTCATTTTA 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1317 TGAACATCAAAAGGAGAA-----AAGCTTCTTGCGAAAAATGTCATATGTTA 1361
QY 1348 ACAATGCTGAACATGTTACACCACTGCACTTCACTAGAAATTTCA 1394
Db 1362 AGGATTAAGGCACTGTTTCTCCTAATGCCCCCTTAATTAATGCTGA 1408

RESULT 3
US-10-437-963-24608
; Sequence 24608, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: LA Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bouhazuk, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 24608
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Oryza sativa
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29576C.1
US-10-437-963-24608
Query Match 22.9%; Score 375; DB 17; Length 1359;
Best Local Similarity 60.0%; Pred. No. 7,66-63;
Matches 646; Conservative 0; Mismatches 425; Indels 6; Gaps 1;

QY 221 AAATATTGAAGACACATATATATTGACAAAGTTGATTAATAATGAGATTAAGATTTA 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 AACTAGTAGAGATGATAGTCTCAGAGGACAGTCTCTTGAAGTCAAAAGAAAGTGAGTTGGA 161
QY 281 TGTACTTAGCTTTGAGACCTAAGGCTGATGGAAGAAACATATGATATATATGCAATTTGAGCA 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 TGTGCGCAACATGCGGCTTACGAGATGACACATGATGACACGAAAGGATTTGGCAAA 221
QY 341 AGCATGGAATGAAGCATGTTTCACTATAGAACACCTGTTCAAATTTGTGTGTTCTTAACAA 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 GGCATGCGGCTGACAGCTTGTCTCTCTTGCAACCTTCATTTGTTCTCATCCCAAGGCGCA 281
QY 401 GAATTAATCTCTGAGCAATTCACCTTTCAAGTCCATGCAAGATCTTCTATTTTCAATPAA 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 GAGATACCTCAACCAAGCACATTAACATCTGTGTCATGCAAAATCAAGCATCACTTCAT 341
QY 461 GATTTTGGATCCCTTAGAAGCATCTAGTAATAATTTTCAGCT-----ACAAAGTATAGAG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 342 GATAGAGGGTACGTTGGTGGCTCTCCAAAGAGTGATGATTTGAGCAAGAAACTATTAG 401
QY 515 GCTTTGATTTGCTTTTGAATGTTCAAAATTTAGTTTGGAGAGAGAACTATCAAA 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 402 GCACGTGAATTAATGTCATATGTTGATGAGTGGCTTACTGTGCTGTGTGTGTGGATGTGCA 461
QY 575 TGGCAATGACAAAGTATGTTGGCCAAAGTTTTCGAAATTAATTAATCACTGCCATGAC 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 462 TGAAGACGCAAAATTTGTGGCAAAATTTCTGCAAAACCAATGCAAACTTCATGACAC 521
QY 635 GGATGACCAACGCGCTTAACCTCTGGAATTCGAAATTTGAAAGTGAATAATCTAAA 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 522 TGAAGCTCAACGCGCTTGAATCTTACCTCTTTCAAACTTAAGTGAAGAACTTGAA 581
QY 695 GAGTAAAAATGACAAACAAATTCATCAAAATTTGAATGATGACATTAATTTGATGCTTC 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 GCTACTAAACAGCAACAAATTCACATGTCAGTTGAGATTTGACATGATTTGAGATCTC 641
QY 755 AAATTTGATGATCAATGCTTCAAGCAAGAGCCCAATTAATGATGATGCTCATGATCAAA 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 TGCCCTGAACATCAACGACAGGACATGACCCCAACACCGAGCATCATATCAAG 701
QY 815 TACTCAATATATTTCAAAATATCTGATCTATTAATTTGGAACAGGATGATTTGATTCAT 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 702 AGGTAAAAATTTCAAGATGACAGGCTGCATTAATCAAGACCGGGATGATGCTGTGAT 761
QY 875 TGTTCGTGATTCAAAATGTCAGGCCCAAAATATTACTTGTGTCAGGTCATGAT 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 762 TGAGGATGAACTGAGAACTTACATGTCAGAAACATGTTGTGTGACCGGGAACAGGCAT 821
QY 935 AAGTATTTGGAAGCTTGAAGATCTGGAATTTCAAGAGCTTATGTCTTAATGTTAATCTPAAA 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 822 CAGCATCGGAGCTTGAAGGATCAACAATCTTAAGCTCATGTCAACAAATGTCACCTGTCA 881
QY 995 TGAAGCCAAATTAATCGGTGCCGAAATGGAATGGAATGAGATCAAGATCTTGACAGGAGATC 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 882 CACCGTCAGGCTATATGAGCAACACACGAGCTGCATCAAGACATGCAAGGCGGTTG 941
QY 1055 TGGACAGCTAGCAACATCAAAATTTCTGAATGTGAATGCAAGACGTTAATATCCAT 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 942 GGGCTACGCGAAGAACATCGTGTTCAGAAACATGATCAAGAAATGTTTGAACCAAT 1001
QY 1115 AATTATAGCAAAACTTTGTGATCGAGTTGAACCATGTATACACAGATTTTCAGAGAT 1174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1002 CATTAATCAACGAACATCTGTGATTTGCTACACCTTCAGAGAAACACATCTGGCGGT 1061
QY 1175 TCAAGTGAATAATGTGTATGAGATATCAAGGACCAAGTGCACAAAGGTGGCAT 1234
```

Db 1062 GCAGGTAGCAATGTGTCTTCAAGAACATCAGGGGACAAAGTGCTTCAAGAGGCTAT 1121
QY 1235 AAAATTGATTGACGACCAAACTTTCCATGTGAAGAAATTATATGAGAAATATATA 1291
Db 1122 CAAACTGACTGTGACGAAATGTACTTGTCAAGAAATTAACCTTGAAGACGTCAA 1178

RESULT 4
US-10-437-963-24607
; Sequence 24607, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 24607
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29575C.1
US-10-437-963-24607

Query Match 22.6% Score 369; DB 17; Length 1182;
Best Local Similarity 60.1%; Pred. No. 1e-61;
Matches 634; Conservative 0; Mismatches 415; Indels 6; Gaps 1;

QY 263 TGGGATTAAAGTGAATTAATGTAATTAAGCTTTGAGAGCTAAGGGTGATGAGAAAACATATGA 322
Db 78 TGGCAGTAATGTGTTCACCATACAGAGCTACGGGGCTCATGAGACGACGATGATGA 137
QY 323 TAATATTCATTTGACGACATGAGAAATGAGCATGTTCATCTTGAACACCTGTCATNT 382
Db 138 CACCAAAACATTTGGAGATGATGAGCTGACAGCTTGCCTCTGCAAAAACCTGACGTTT 197
QY 383 TGTGTTCTTAAACCAAGAAATATCTTCTCAAGCAATCACTTTGAGTCCATGAG 442
Db 198 GCTATCCCAAGGCAAGAAATACCTGATCAAGCAACACACATGTCTGTCATGCA 257
QY 443 ATCTTCTATTCAGTAAGATTTTGTGATCTCTCAAGCAATCACTTTGAGTCCATGACT- 501
Db 258 ATCAAGCATCTATGATGTGTGAGGGTATGTTGGTGGCTTCTCCAGAGAGGTCAAGCTG 317
QY 502 -----ACAAAGTGAAGGCTTTGATGCTTTGATGATGTTCAAAATTAGTGTGG 556
Db 318 GAGCAAGGAGACCATTAAGGACCTGATTCGATCAAGTGTGATCACTGCTTACTGTGAC 377
QY 557 AGAGAGAGGAATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 616
Db 378 TGGTGTGGGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 437
QY 617 TAAATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
Db 438 CTCGAAGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
QY 677 GAAAGTGAATTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 736
Db 498 GAAGGTAGAGTATCTAAGGTGTGAAAGCCAGCAAAATTCAGATTTCAATGAGGATTTG 557
QY 737 CACTATGTTGATGCTTCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 796

Db 558 CACCATGTGATGTGTCTGCTGCTGATCAGACACGAAACAGCCCAACTGA 617
QY 797 TGGAGTCCATGTAATCAATCTCAATATATTTCAATATCTGATATTTATTTGAAACAG 856
Db 618 TGAATTCATATCAACGCGACGACAGATGTTGAGTACAGATCTGATATCAAGACGG 677
QY 857 TGATGATTTGATTTCAATTTTCTGATCTCAAAATGTGAGGCGCAAAATTTACTTG 916
Db 678 GGATGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
QY 917 TGGTCAAGTCAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 976
Db 738 TGGACCGGACACGCGCATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 797
QY 977 GTCTAATGTAAGTAAATGAAGCCAAATTAATGATGATGATGATGATGATGATGATGAT 1036
Db 798 CAACATGTCACTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
QY 1037 GACTTGGCAGGAGATCTGACAAAGCTTACACATCAATTTGTAATGTGAATGCA 1096
Db 858 GACATGCGAGGAGGAAAGGCTTACGAGAAAGATGATGATGATGATGATGATGATGATGAT 917
QY 1097 AGAGTTAATGATCCCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1156
Db 918 CAATGTCTGAACCCATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 977
QY 1157 ACACAGTTTTCAGCAGTTCAGTGAATGAATGATGATGATGATGATGATGATGATGATGAT 1216
Db 978 GCAACAGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1037
QY 1217 TGCAACAAAGGTGCGCATTAATTTGATGATGATGATGATGATGATGATGATGATGAT 1276
Db 1038 TGATCAAGAGAGGACCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1097
QY 1277 AATGAGAAATTAATTTAGTGAAGGAAATGAA 1311
Db 1098 CTGGAAGATGATCATCTCACTGTCAGAGGAGAA 1132

RESULT 5
US-09-782-130-26
; Sequence 26, Application US/09782130
; Publication No. US20040055038A1
; GENERAL INFORMATION:
; APPLICANT: KNAUF, VIC C.
; APPLICANT: KNAUF, VIC C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR REGULATED TRANSCRIPTION
; FILE REFERENCE: 16518.052
; CURRENT APPLICATION NUMBER: US/09/782,130
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: US 09/232,861
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: US 08/812,665
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 08/484,941
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/105,852
; PRIOR FILING DATE: 1993-08-10
; PRIOR APPLICATION NUMBER: US 07/526,123
; PRIOR FILING DATE: 1990-05-21
; PRIOR APPLICATION NUMBER: US 07/267,685
; PRIOR FILING DATE: 1988-11-02
; PRIOR APPLICATION NUMBER: US 06/592,605
; PRIOR FILING DATE: 1985-01-17
; PRIOR APPLICATION NUMBER: US 07/582,241
; PRIOR FILING DATE: 1990-09-14
; PRIOR APPLICATION NUMBER: US 07/188,361
; PRIOR FILING DATE: 1988-04-29
; PRIOR APPLICATION NUMBER: US 07/168,190
; PRIOR FILING DATE: 1986-03-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 35


```

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30261
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73002C03_FLI
US-10-425-114-30261
```

```

Query Match      16.3%; Score 267.4; DB 13; Length 1308;
Best Local Similarity 58.7%; Pred. No. 5.9e-42;
Matches 463; Conservative 0; Mismatches 326; Indels 0; Gaps 0;
```

```

QY 500 CTACAAAGATGAAAGGCTTGGATTGCTTTTGATAGTGTCAAAATTTAGTTGTGGAGG 559
DB 64 CGAACATGACGAGGACATGATCGTGTCCGAGCATGACAGCTCACCGCACAACG 123
QY 560 AGGAGGAATATCAATGAGCAATGACAAAGTATGTTGGCCAAATTCTTGCAAAATAA 619
DB 124 CGGTGGCGCATGATGAGCAACGCGGAAATGTTGGCTTCACCTCGTGAAGATCAACAA 183
QY 620 ATCACTGCCATGACGAGGATGACCAACGCGCTTAACCTTGGAAATGCAAAATTTGAA 679
DB 184 GGCCTCCCTTTGCAAGAGGCTCCGAGCTCTGTCTATTCACCTGCGTGCACCTGAA 243
QY 680 AGTGAATATCTAAAGATGAAAAATGCAACAACAAATTCATATCAAAATTTGAGTCATG 729
DB 244 AGTGGAGATCTGAAATCGTGAACAGCCAAAGATACATGTGATGAGGATTTGTC 303
QY 740 TAATGTGATGCTCAATTTGATGATCAATGCTTACGCAAAAGCCCAATATCTGATG 799
DB 304 AAATGCGCTGCTGCGCGGTGTTCATCAACGCGCTGCGACCGCTTAACCTGACG 363
QY 800 AGTCACTGATCAATATCTCAATATATTAATCTGATGATCTATTTAGTGAACAGTGA 859
DB 364 CATCCATCATCGCGAGCAAGAAATGTACGCTCAACAGCTGCAAGATCAAGCAGGGA 423
QY 860 TGATTTGATTTCAATTTGTTCTGGATCTCAAAATGTCAGGCGCAAAATTTACTTGTG 919
DB 424 CCACTGCAATGATGAGAAACGCGGACTCAACCTTCATGTCTCAAAAGTTAATCTGTG 483
QY 920 TCCAGGTCAATGATGATGATGAGAAAGCTTGAATCTGAAATTCGAAGCTTATGTC 979
DB 484 TCCAGGCGCATGAGATGAGATCGGAAGCCTAGAGACGAACAATCAAGAGCAGAACTTC 543
QY 980 TAATGTTACTGTAATGAAGCCAAATTAATCGGTGCGGAATGAGGTTAGATCAAGC 1039
DB 544 CGGCATCAACCATAGATTCAGTGTCACTACACGCAACCAACGAGGACGCAATCAAGC 603
QY 1040 TTGGCAAGGAGGATCTGCAACAGCTAGCAACATCAAAATTTGAAATGTGAAATGCA 1099
DB 604 GTACACAGGAGGACAGGATACGCCAAGGACATCAAGTTCAAAATATGTTATGTACGA 663
QY 1100 GCTTAAGATCCCATTAATTAATGACCAAAATTTGATGATGAGTTGAACATGTATACA 1159
DB 664 CATTCAAGAACCCGATATATATGACCAAGACTACGCGCAAGGCTTAAGCCATGCGGAGA 723
QY 1160 ACAAGTTTCAGAGTTCAAGTGAATAATGTGTTATGAGATATCAAGGGGACAAAGTC 1219
DB 724 ACAAGATGACAGTGCAGGTGACGGGTGTCTTTCAGAAACATTTAGAGGACACAG 783
QY 1220 AACAAAGGTGCGCATTAATTTGATTTGACACAAACTTTCCATGTAAGAAATTAAT 1279
DB 784 TACCAAGAGACGCATCAGATGAACTGCAAGTGAAGACGTCCATCCAAAGGATTACTT 843
QY 1280 GGAAGATAT 1288
DB 844 GCAAGACAT 852
```

RESULT 9

```

US-10-437-963-24606
; Sequence 24606, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 24606
; LENGTH: 1809
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29574C.1
US-10-437-963-24606
```

```

Query Match      14.1%; Score 230; DB 17; Length 1809;
Best Local Similarity 60.0%; Pred. No. 1.3e-34;
Matches 383; Conservative 0; Mismatches 255; Indels 0; Gaps 0;
```

```

QY 657 TTCTGAATTTGCAAAATTTGAAAGTGAATATCTAAAGGTTAAATGCAACAATTT 716
DB 142 TGCAGAGACATTTTATTTCTGAAGTGAACATCTGAGGTGTTGAACGCCGCAATC 201
QY 717 CATATCAAAATTTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 776
DB 202 CAGATTTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 261
QY 777 GCANAAGCCCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 836
DB 262 GAAATCTGCCCCAAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 321
QY 837 GATATATTTATTTGAAACAGTATGATGATGATGATGATGATGATGATGATGATGATG 896
DB 322 GATGACAGATCAAGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 381
QY 897 CAGGCCAACAATTAATCTTGTGTCAGGTCAATGATGATGATGATGATGATGATGATG 956
DB 382 CATGTCAGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 441
QY 957 GGAATTTGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1016
DB 442 CATTAATTTGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
QY 1017 GAAATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1076
DB 502 AGCATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
QY 1077 TTTGTAATTTGGAATGCAAGACCTTAAGTATCCCATTAATTTAGAACCAAAATTA 1136
DB 562 TTCCAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
QY 1137 GATGAGTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1196
DB 622 GATCTTTTACACCTTGAAGACCAAGAAATCTGAGTGAAGTGAAGTGAAGTGAAGTGA 681
QY 1197 GAGATATCAAGGACAGTGAACCAAGGTCGATTAATTTGATGACGACCAAC 1256
DB 682 AAGAACATCAAGGAGCAAGTGAACATGATGATGATGATGATGATGATGATGATGATG 741
QY 1257 TTTCATGTTGAAGATTAATTAATGATGATGATGATGATGATGATGATGATGATG 1294
DB 742 GTGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779
```

RESULT 10
US-10-424-599-61964
; Sequence 61964, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 61964
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26969C.1
US-10-424-599-61964

Query Match 12.9%; Score 210.8; DB 13; Length 1333;
Best Local Similarity 54.0%; Pred. No. 5.9e-31;
Matches 431; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
DB 512 AAGCTTTGATGCTTTGATGATGTCATTAATTTAGTTGAGAGAGAGAGAACTAT 571
19 ACGGTTTGGCTTATTTTCCAAATGACAAACATTTTCCAGGTCCTGAGTTAT 78
QY 572 CAATGGCAATGACAGATGAGTGGCCCAAGTTCTTGCAAAATTAATTAATCACTGCCATG 631
79 TGATGGCTCAGAGAGCAAAATGATGGCGCATCTTGGCAAAAAGAACAAAGTCCATCTTG 138
QY 632 CAGGATGACCAACGCGCTTAACCTTGGAATTTGCAAAATTTGAAGTGAATATCT 691
139 CAAGGTGACCAACAGATTTACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 198
DB 692 AAAGATTAATGACCAACAAATTCATATCAAAATTTAGTCAGTCACTAATGTTGAGC 751
199 AACATCCAGATGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 258
QY 752 TTCAATTTGATGATCAATGCTTTCAGCAAAAGGCCCAATTAATCTGATGATGATGATGATGAT 811
259 TACTGGCGTGAAGTGTGACGACCTGAGACAGCCCAACACTGATGATGATGATGATGATGAT 318
DB 812 AAATATCTCAATATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAAT 871
319 TGAATCAACAAATGATCAATGATCAACAGCAAAATTTGAGACAGGAGATGATGATGATGAT 378
QY 872 AATTGTTCTGATGATCAAAATGATGAGGCCCAAAATATTAATTTGATGATGATGATGATGAT 931
379 AATTGTCATGCTAGCTCAATATCAAAATGAGAAATTTATTTGAGACAGGACATG 438
DB 932 TATAAGTATGAGAGCTTNGATCTGGAATTCAGAACTGATGATGATGATGATGATGATGATGAT 991
439 AATCAGCATTTGAGAGCTTNGAGAGAAAGCACTCAACAGGATGATGATGATGATGATGATGAT 498
QY 992 AAATGAGCAAAATTAATGATGAGGCCCAAAATGAGTGAATCAAGATTTGAGAGGAG 1051
499 GGAATCAGAGATTTTGAAGGAGCTACCAACAGGCTCAGAAATTAAGATTTGAGAGGAG 558
DB 1052 ATCTGAGCAAGCTAGCAACATCAAAATTTCTGATGATGAGAAATGCAAGCTTGAATGCC 1111
559 TTCTGATATGTTTGAAGGAGGCTTTCAGATGATGAGAGGAGGAGAAATTCAGAGCGAT 618
QY 1112 CATTAATTAAGCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171
619 CGGCTCGAGAGCAATTTTACTGATGATTTCTCCACCAATTTGAAATTAAGGATCAGC 678

QY 1172 AGTTCAAGTGAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
DB 679 AGTGAGATGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 738
QY 1232 CATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
DB 739 CATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 798
QY 1292 TTTAGTAGGGAAGTGG 1309
DB 799 CTTAGAGAAACAGATGG 816

RESULT 11
US-10-425-114-10970
; Sequence 10970, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 10970
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 70092481_F11
US-10-425-114-10970

Query Match 12.1%; Score 198; DB 13; Length 1219;
Best Local Similarity 55.1%; Pred. No. 1.7e-28;
Matches 387; Conservative 0; Mismatches 315; Indels 0; Gaps 0;
DB 608 CAAATTAATTAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 667
1 CAAAAGAACAGTCAATCTTGGCAAGGTGACCAACAGATTTACATGATGATGATGATGATGAT 60
QY 668 CAAAATTTGAAAGTGAATTAATCAAGATGAATAAATGACAAACAAATTCATATCAAT 727
DB 61 TTCAATCCATAGGAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 728 TGAT 787
DB 121 ATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 788 AAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
DB 181 AAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 848 TGAACAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
DB 241 TGAACAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 908 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967
DB 301 AATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 968 AGCTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
DB 361 AGGAT 420
QY 1028 TAGGATCAAGATTTGAGAGGAGATCTGAGACAAAGTGAACATCAATTTCTGAATGT 1087
DB 421 CAGATTAAGATTTGAGAGGAGGATCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY 1088 GGAAATGCAAGCTTAAGTATCCCATTAATATAGACCAAACTATTGTGATCGAGTTGA 1147
| | | | |
Db 481 GAGGCTGAAATATATATCCAACTTATTAAGACCAATTTACTGTGATTCCTCAAC 540
| | | | |
QY 1148 ACCATGATATCAACAGTTTTCAGCAGTTTCAAGTAAATGTGTGATAGATATCA 1207
| | | | |
Db 541 CAATTGGAAATATGAGCATTCAGATGAGATGAGATGAGATGATGATGAGATGAG 600
| | | | |
QY 1208 GGGCAACAATGTCACAAAGGTGGCCATTAATTTGATTTGATGAGCAACAACTTTCATGTA 1267
| | | | |
Db 601 TGGCAGTACCAATAGTGTAAAGGCTTAATTTGATGATGAGATGATGATGATGAG 660
| | | | |
QY 1268 AGGAATTAATATGAGATTAATTAATTTGATGAGGAAAGTGG 1309
| | | | |
Db 661 CAAACTAGTCTTACGACGTTGATTAAGAGAAACAAGATGG 702
| | | | |

RESULT 12

US-10-424-599-141807
; Sequence 141807, Application US/10424599
; Publication No. US2004031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141807
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_99063C.1
US-10-424-599-141807

Query Match 11.9%; Score 194.8; DB 13; Length 1479;
Best Local Similarity 53.5%; Pred. No. 8e-28; Indels 33; Gaps 4;

Matches 518; Conservative 0; Mismatches 417; Indels 33; Gaps 4;
QY 273 GTGATTAATGATCTTACCTTGAAGCTAAGGTGATGAAAAACATATGATTAATTTGA 332
| | | | |
Db 286 GTTTTCGATGTGAGATCTTTGGGGCTGTGAGATGGCTGTGATGATGACACAGTGA 345
| | | | |
QY 333 TTTGAGCAAGCATGGAATGAAAGCATGTTCTAGAACACCTGTTCAATTTGTGTTCT 392
| | | | |
Db 346 TTCAAGCAAGCATGGAAGCAAGCATGT--GCTGTGACTCTGGGATTTGTTCTCTCA 402
| | | | |
QY 393 AAAAACAAGATTTATCTTCTCAAGCAAAATCACCTTTTCAAGTCATGAGATCTTCTAT 452
| | | | |
Db 403 GAAAACCTACAGTTTAAATCACTTCACTATTTTTCAGGTCTCAAGCAAGCATG 462
| | | | |
QY 453 TCAGTAAGATTTTGTGATCTT-----AGAGCATCTAGTAATTAATTTCAAGTAC 503
| | | | |
Db 463 GTATTCAGATGATGAGAACTAATGGAACCAAGTGAACCAATTCGTGGCCAAAGCA 522
| | | | |
QY 504 AAAAGTAAAGGCTTTGATGCTTTGATGATGTTCAAAATTTAGTTGTGAGAGAGA 563
| | | | |
Db 523 GATAGCCCAATCAATGCTGTGATTTATGCACTTGAACCAATGACTTTAATGTGACA 582
| | | | |
QY 564 GAAACTATATGCAATGAGCAAAAGTATGTTG-----CCAACTTCT 605
| | | | |
Db 583 GGAACCATTAAGGAAATGAGAAACAAATGTTGGATCTTCCCTGCAAGCCTTCAAGGGGT 642
| | | | |
QY 606 TGCAAAATTAATTAATCACTGCAATGCAAGGATGCAACCAAGCCTTCACTTGTGAAT 665
| | | | |
Db 643 CCCAATGAAAAAATTTGTCAAGGCAATGTGTGATCTCTGATGATGAGGTTCTTCAATG 702
| | | | |

QY 666 TGCAAAATTTGAAGTGAATTAATCTAAAGATTAATGAACCAAAATTCATATCAAA 725
| | | | |
Db 703 AGCTCCAAATTTGAAGTGAAGGGGTTGAATTAATTAACAACTCTCAGTTCCATGATA 762
| | | | |
QY 726 TTTGATGATGACATTAATTTGATGCTTCAAAATTTGATGATCAATGCTTCAAGAAAGC 785
| | | | |
Db 763 TTCAATGCTGCCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 822
| | | | |
QY 786 CCAATATCTGATGAGTCCATGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 845
| | | | |
Db 823 CCAACACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 882
| | | | |
QY 846 ATTGAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 905
| | | | |
Db 883 ATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 942
| | | | |
QY 906 AATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 965
| | | | |
Db 943 GGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1002
| | | | |
QY 966 GAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1025
| | | | |
Db 1003 CAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1062
| | | | |
QY 1026 GTTAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1085
| | | | |
Db 1063 CTCAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1122
| | | | |
QY 1086 GTGAAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1145
| | | | |
Db 1123 ATCCAAATGAGAACTTGAACCTGAAACCTGATCATATATAGCAATATGATGATGATGATG 1182
| | | | |
QY 1146 GAACCATGATTAACAAGTTTTCAGAGATGATGATGATGATGATGATGATGATGATGATG 1205
| | | | |
Db 1183 GA--ATGCTTAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
| | | | |
QY 1206 AAGGCAAC 1213
| | | | |
Db 1240 AAGGCAAC 1247
| | | | |

RESULT 13

US-09-770-444-874/C
; Sequence 874, Application US/09770444
; Patent No. US20020023280A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krieker, Ted
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Huban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2027 (PABA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; NUMBER OF SEQ ID NOS: 999

SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 874
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-444-874

Query Match 10.7%; Score 175; DB 9; Length 442;
Best Local Similarity 65.0%; Pred. No. 3.1e-24; Indels 3; Gaps 1;
Matches 275; Conservative 0; Mismatches 145;

QY 963 TCAGAAAGCTTATGCTTAATGTTACTGTAATGAAGCCAAAATTATCGTCCGAAAT 1022
DB 441 TCAGAAAGCATATGTTCAATGTTGTTAGTCAACAAAGCACTCTTATAGGAACCACTAAT 382
QY 1023 GAGATTAGATCAAGACTTGCGAGGAGGATCTGACAACTAGCAACATCAATTTCTG 1082
DB 381 GGTGTGAGAAATCAAGACTTGCGAGGAGGACATGGAATGGCAAGAACATCATATTTCCAA 322
QY 1083 AATGTGAAATGCAAGCGTTAAGTATCCATTAATTAGCCAAACTATTGTATCGA 1142
DB 321 GACATCATATGAAAACGTTACAAACCCCAATTAATCATCAACGAGCTACTGGATCGT 262
QY 1143 GTTGAACCATGTATACAAAGTTTTCAGCACTTCAGTGAAGAAATGTGTGTATGAGAA 1202
DB 261 GTTGAAGCATGCCCCGAAACAGAAATCCGCGGTACAGTGAAGCAATGTGTGTACAAAAC 202
QY 1203 ATCAAGGCGCAAGTGCACAAAGGTGCGCATTAATTTGATTGACGACAACTTTCCA 1262
DB 201 ATACAAAGGACGAGCTCAAGACCCCATAGCTGTGAATTTGTATCAGCAAGAACATTTCCA 142
QY 1263 TGTGAAGAAATTAATAGGAAATATTAATTTAGAGGGAAGTGGAAAA---CCATCA 1319
DB 141 TGTGAGAAATATCAATGCAAAAGCTTAATCTGTCAATCAACCAAGATGATCC 82
QY 1320 GAGGCTACGTGCAAAATGTCATTTTAACAATCTGAAACATGTTACACCACTGCACT 1379
DB 81 AAAGCTTCTTGCTCCATGTGAAGTTGAGACACCGAGAAATGTTTCTCTTTGCACT 22
QY 1380 TCA 1382
DB 21 TGA 19

RESULT 14
US-09-924-035A-369/c
; Sequence 369, Application US/09224035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grilach, Jrm
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924.035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 369
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-924-035A-369

Query Match 10.7%; Score 175; DB 9; Length 442;
Best Local Similarity 65.0%; Pred. No. 3.1e-24;
Matches 275; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 963 TCAGAAAGCTTATGCTTAATGTTACTGTAATGAAGCCAAAATTATCGTCCGAAAT 1022
DB 441 TCAGAAAGCATATGTTCAATGTTGTTAGTCAACAAAGCACTCTTATAGGAACCACTAAT 382

QY 1023 GAGATTAGATCAAGACTTGCGAGGAGATCTGACAAAGCTTACCAATCAATTTCTG 1082
DB 381 GGTGTGAGAAATCAAGACTTGCGAGGAGGACATGGAATGGCAAGAACATCATATTTCCAA 322
QY 1083 AATGTGAAATGCAAGCGTTAAGTATCCATTAATTAGCCAAACTATTGTATCGA 1142
DB 321 GACATCATATGAAAACGTTACAAACCCCAATTAATCATCAACGAGCTACTGGATCGT 262
QY 1143 GTTGAACCATGTATACAAAGTTTTCAGCACTTCAGTGAAGAAATGTGTGTATGAGAA 1202
DB 261 GTTGAAGCATGCCCGAAACAGAAATCCGCGGTACAGTGAAGCAATGTGTGTACAAAAC 202
QY 1203 ATCAAGGCGCAAGTGCACAAAGGTGCGCATTAATTTGATTGACGACAACTTTCCA 1262
DB 201 ATACAAAGGACGAGCTCAAGACCCCATAGCTGTGAATTTGTATGACCAAGAACATTTCCA 142
QY 1263 TGTGAAGAAATTAATAGGAAATATTAATTTAGTAGGGAAGTGGAAAA---CCATCA 1319
DB 141 TGTGAGAAATATCAATGCAAAAGCTTAATCTGTCAATCAACCAAGATGATCC 82
QY 1320 GAGGCTACGTGCAAAATGTCATTTTAACAATGCTGAACATGTTACCACTGCACT 1379
DB 81 AAAGCTTCTTGCTCCATGTGAAGTTGAGACACCGAGAAATGTTTCTCTTTGCACT 22
QY 1380 TCA 1382
DB 21 TGA 19

RESULT 15
US-10-362-091-3
; Sequence 3, Application US/10362091
; Publication No. US20040049809A1
; GENERAL INFORMATION:
; APPLICANT: Instituto de Ciencia Aplicada e Tecnologia (ICAT)
; TITLE OF INVENTION: Pear genes codifying for b-galactosidase, Pectin Methyl-esterase,
; TITLE OF INVENTION: Polygalacturonase, Expansin and their use.
; FILE REFERENCE: none
; CURRENT APPLICATION NUMBER: US/10/362.091
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: PT 102511 C
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Pyrus communis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(1308)
; OTHER INFORMATION:
US-10-362-091-3

Query Match 10.5%; Score 171.2; DB 13; Length 1673;
Best Local Similarity 54.9%; Pred. No. 3.2e-23;
Matches 360; Conservative 0; Mismatches 293; Indels 3; Gaps 1;

QY 639 GCACCAAGCGCTTAACCTTCTGAAATGCAAAATTTGAAAGTAATCTAAGAGT 698
DB 562 GAGGCACTACACTAGAGCTTTTCGAACCTCCAAACAGTTGTGTAGTGAATTAATATCA 621
QY 639 AAAAAATGCAACAATAATCATATCAAAATTTGATGATGACTAATAGTTGTACTTCAAT 758
DB 622 CTAAACAGCAAAATGTTCCACTTGTGCTCAACGCGCTGCCAAAATGTGAAATGCAAGGT 681
QY 759 TTGATGATCAATGCTTTCAGCAAGAGCCCAATTAATGAGTGGATGATATCAATATCT 818
DB 682 GTCAAGGTTAAAGCGCGCGGCAACAGCCCAACACCGAGGATCATCATCAAAATGCTA 741
QY 819 CAATATATTCAATATCTGATCTATTTATGGAACAGTGAATGATTTTCAATGTT 878
DB 742 TCTGAGATCACTTCTGACTTCCAAATTTTCAACCGGTGAGGACGTGTCTCAGTTGGC 801

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 26, 2004, 23:07:30 ; Search time 4442 Seconds

(without alignments)
10998.318 Million cell updates/sec

Title: US-10-018-604-1

Perfect score: 1636
Sequence: 1 aaacttttcaatgacaag.....atatgaaaaaaaaaaaaaa 1636

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_huv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdg:*
26: em_gss_pmg:*
27: em_gss_vr1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778.6	47.6	797	12	BM410796 EST585123
2	770	47.1	786	12	BM412817 EST587144
3	765.4	46.8	767	12	BM412528 EST586855
4	754	46.1	774	12	BM413004 EST587331

5	742	45.4	742	12	BM413282	BM413282	EST587609
6	738	45.1	748	12	BM413343	BM413343	EST587670
7	726.4	44.4	729	12	BM413075	BM413075	EST587402
8	709	43.3	729	10	AM222257	AM222257	EST299068
9	707.4	43.2	709	10	AM441939	AM441939	EST311335
10	694.4	42.4	696	10	AM442253	AM442253	EST311649
11	683.8	41.8	687	10	AM442335	AM442335	EST311731
12	674	41.2	703	12	BM536089	BM536089	EST589111
13	673.8	41.0	677	10	AM442241	AM442241	EST311637
14	670.2	41.0	675	12	BM408546	BM408546	EST582873
15	657.2	40.2	670	12	BM411547	BM411547	EST585874
16	657	40.2	672	10	AM222222	AM222222	EST299033
17	654.4	40.0	672	10	AM221785	AM221785	EST298596
18	643	39.3	643	10	AM223400	AM223400	EST300211
19	642.4	39.3	644	10	AM442052	AM442052	EST311448
20	639	39.1	639	10	AM442280	AM442280	EST311676
21	631	38.6	632	10	BE434493	BE434493	EST405571
22	630.4	38.5	632	12	BM410538	BM410538	EST584865
23	628.8	38.4	632	10	BE432845	BE432845	EST399470
24	628.4	38.4	630	10	BE431613	BE431613	EST336428
25	623	38.1	623	12	BM409252	BM409252	EST583579
26	617.4	37.7	619	12	BM413302	BM413302	EST587529
27	617	37.7	617	10	BE460938	BE460938	EST412357
28	604	36.9	604	10	BE433321	BE433321	EST39850
29	603	36.9	614	10	BE433180	BE433180	EST39709
30	602.4	36.8	604	10	BE461204	BE461204	EST412623
31	597.8	36.5	601	10	AM222835	AM222835	EST299646
32	595.4	36.4	598	10	AM223911	AM223911	EST300722
33	588	35.9	600	10	AM442198	AM442198	EST311594
34	583.4	35.7	585	10	BE436917	BE436917	EST408035
35	583.4	35.7	593	10	AM222726	AM222726	EST299537
36	580	35.5	600	10	AM441498	AM441498	EST310894
37	579	35.4	582	12	BM409146	BM409146	EST583473
38	574.4	35.1	576	10	AM222864	AM222864	EST299675
39	572	35.0	572	10	AM223468	AM223468	EST300279
40	570.4	34.9	572	10	AM222731	AM222731	EST299542
41	570.4	34.9	588	10	BE432533	BE432533	EST399062
42	567	34.7	567	10	AM223661	AM223661	EST300472
43	565	34.5	566	10	BE434462	BE434462	EST405540
44	563	34.4	563	10	AM441402	AM441402	EST310798
45	561.8	34.3	565	10	BE460910	BE460910	EST412329

ALIGNMENTS

RESULT 1
LOCUS BM410796 797 bp mRNA linear EST 22-JAN-2002
DEFINITION EST585123 tomato breaker fruit Lycopersicon esculentum cDNA clone
ACCESSION CLE65415 5' end, mRNA sequence.
VERSION BM410796
KEYWORDS BM410796.1 GI:18262426
SOURCE EST.
ORGANISM Lycopersicon esculentum (tomato)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 797)
Alcala,J., Vrebalov,J., White,R., Vision,T., Karanycheva,S.A., Tsai,J., Bougri,O., Kirtnes,E., Uteback,T., Van Aken,S., Roming,C.M., Fraser,C.M., Martin,G.B., Tankeley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue, breaker stage (2002)
Unpublished (2002)
Contact: CGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics Institute

TITLE JOURNAL COMMENT
BM410796 EST585123
BM412817 EST587144
BM412528 EST586855
BM413004 EST587331

FEATURES Seq primer: T3.
Location/Qualifiers
Source 1..797
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG54L15"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCldapt; Site 1: EcoRI;
/clone_1lb="tomato breaker fruit"
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

Query Match 47.6%; Score 778.6; DB 12; Length 797;
Best Local Similarity 99.4%; Pred. No. 2.3e-120;
Matches 792; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 359 TTCAATGAGACACCTGTTCAATTGTTGCTCAAAAACAAGAAATATCTTCTCAAGCA 418
DB 1 TTCAATGAGACACCTGTTCAATTGTTGCTCAAAAACAAGAAATATCTTCTCAAGCA 60
QY 419 AATGACCTTTTCAGATCCATGACATCTTCTATTTCAGTAAGAATTTTGGATCCTTGA 478
DB 61 AATGACCTTTTCAGATCCATGACATCTTCTATTTCAGTAAGAATTTTGGATCCTTGA 120
QY 479 AGCATCTAGTAATAATTTCAGACTACAAAGATGAAAGCTTTGGATGCTTTTGATAGT 538
DB 121 AGCATCTAGTAATAATTTCAGACTACAAAGATGAAAGCTTTGGATGCTTTTGATAGT 180
QY 539 TCAGAAATTTAGTTTGGAGAGAGAGAACTATCAATGAGCAATGAGCAAGTATGGTGGCC 598
DB 181 TCAGAAATTTAGTTTGGAGAGAGAGAACTATCAATGAGCAATGAGCAAGTATGGTGGCC 240
QY 599 AAGTCTTGGCAAAATTAATAATCACTGCCATGACAGGATGACCAAGGCTTAACTT 658
DB 241 AAGTCTTGGCAAAATTAATAATCACTGCCATGACAGGATGACCAAGGCTTAACTT 300
QY 659 CTGGAATTTGCAAAATTTGAAAGTGAATATCTTAAAGCTTAAATGACACAAATTTCA 718
DB 301 CTGGAATTTGCAAAATTTGAAAGTGAATATCTTAAAGCTTAAATGACACAAATTTCA 360
QY 719 TATCAAAATTTAGTCATGACATTAATGTTGAGCTTCAAAATTTGATGATCATGCTTCAGC 778
DB 361 TATCAAAATTTAGTCATGACATTAATGTTGAGCTTCAAAATTTGATGATCATGCTTCAGC 420
QY 779 AAAGAGCCCAATATCTGATGAGTCATGATCAAAATCTCAATATATTCAAATATCTGA 838
DB 421 AAAGAGCCCAATATCTGATGAGTCATGATCAAAATCTCAATATATTCAAATATCTGA 480
QY 839 TACTATTTATTTGGAACAGGTGATGATTTTCAATTTGTTCTGGATCTCAAAATGTCGA 898
DB 481 TACTATTTATTTGGAACAGGTGATGATTTTCAATTTGTTCTGGATCTCAAAATGTCGA 540
QY 899 GGCCCAAAATTTCTTGTGTCAGGTCATGTTAAGTATTTGAGACTTGAAGATCTGG 958
DB 541 GGCCCAAAATTTCTTGTGTCAGGTCATGTTAAGTATTTGAGACTTGAAGATCTGG 600
QY 959 AAATTCAGAGCTTATGTGTCTAATGTTACTGTAATGAAAGCCAAATTTATCGTGCGCA 1018
DB 601 AAATTCAGAGCTTATGTGTCTAATGTTACTGTAATGAAAGCCAAATTTATCGTGCGCA 660
QY 1019 AAATGAGTTAGATCAAGACTTGGCAGGAGAGATCTTGAGACAGCTAGACATTAATTT 1078
DB 661 AAATGAGTTAGATCAAGACTTGGCAGGAGAGATCTTGAGACAGCTAGACATTAATTT 720

QY 1079 TCTGATGTTGAATGCAAGACGTTAAGTATCCCATATTAATATGACCAAACTATT-GTG 1137
DB 721 TCTGATGTTGAATGCAAGACGTTAAGTATCCCATATTAATATGACCAAACTATTGTG 780
QY 1138 ATCGAGTTGAACCATGT 1154
DB 781 ATCGAGTTGAACCATGT 797

RESULT 2
BM412817 786 bp mRNA linear EST 22-JAN-2002
LOCUS EST587144 tomato breaker fruit Lycopersicon esculentum cDNA clone
cLEG54A20 5' end, mRNA sequence.
BM412817
VERSION BM412817.1 GI:18264447
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 786)
Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A.,
Tsai,J., Bougri,O., Kirkness,E., Uteerback,T., Van Aken,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue, breaker stage (2002)
Unpublished (2002)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
Source 1..786
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG54A20"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_1lb="tomato breaker fruit"
/note="Vector: pBluescriptSKmCldapt; Site 1: EcoRI;
Site 2: XhoI; supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

Query Match 47.1%; Score 770; DB 12; Length 786;
Best Local Similarity 98.7%; Pred. No. 6.3e-119;
Matches 776; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 333 TTGAGCAGATGAGATGAGCATGTTCACTAGAACACCTGTTCAATTTGTGGTTCT 392
DB 1 TTGCGCAGAGATGAGATGAGCATGTTCACTAGAACACCTGTTCAATTTGTGGTTCT 60
QY 393 AAAAAAGAAATATCTTCTCAAGCAATCACTTTTGAAGTCATGACATCTTCTATT 452
DB 61 AAAAAAGAAATATCTTCTCAAGCAATCACTTTTGAAGTCATGACATCTTCTATT 120
QY 453 TCAGTAAGATTTTGGATTCCTTGAAGCATCTAGTAATTTTTCAGACTAGCAAGATAGA 512
DB 121 TCAGTAAGATTTTGGATTCCTTGAAGCATCTAGTAATTTTTCAGACTAGCAAGATAGA 180


```

/lab_host="SOLR"
/clone_lib="tomato breaker fruit"
/notes="Vector: pBluescriptSKcudap; Site 1: EcoRI;
Site 2: XhoI; supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

```

ORIGIN

```

Query Match      45.4%; Score 742; DB 12; Length 742;
Best Local Similarity 100.0%; Pred. No. 36-114;
Matches 742; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 TGGTATCCAAAGAAATAGATATCTCCTCTCATTTATTTTGGTCAATTCOA 111
DB 1 TGGTTATCCAAAGAAATAGATATCTCCTCTCATTTATTTTGGTCAATTCOA 60

QY 112 CTTGTGAAGCAATGTTATGTAGACATTTATTCAAAGTTTATGATATTTCTTG 171
DB 61 CTTGTGAAGCAATGTTATGTAGACATTTATTCAAAGTTTATGATATTTCTTG 120

QY 172 AACAGAAATTTGCTATGATTTTCAAGCTTATCTTTCTTATTTGACAAAATTTGAA 231
DB 121 AACAGAAATTTGCTATGATTTTCAAGCTTATCTTTCTTATTTGACAAAATTTGAA 180

QY 232 GCAACATATATTTGCAAGGTTGATTAATAATGGATTAAGGATTAATGCTAGCT 291
DB 181 GCAACATATATTTGCAAGGTTGATTAATAATGGATTAAGGATTAATGCTAGCT 240

QY 292 TTGGAGCTAAGGCTGATGAAAAAATATGATATATTTGATTTGAGCAAGCATGGAATG 351
DB 241 TTGGAGCTAAGGCTGATGAAAAAATATGATATATTTGATTTGAGCAAGCATGGAATG 300

QY 352 AAGCATGTTCAATCAACACCTGTTCAATTTTGGTCTTAATAACAAGATTTATCTTG 411
DB 301 AAGCATGTTCAATCAACACCTGTTCAATTTTGGTCTTAATAACAAGATTTATCTTG 360

QY 412 TCAGCAAAATCACTTTTCAAGGTCAGTGCAGATCTTCTATTTGATTAAGATTTTGGAT 471
DB 361 TCAGCAAAATCACTTTTCAAGGTCAGTGCAGATCTTCTATTTGATTAAGATTTTGGAT 420

QY 472 CCTTGAAGCATCTAGTAAATTTTCAAGCTACAAAGATGAAAGCTTTGGATGCTTTTG 531
DB 421 CCTTGAAGCATCTAGTAAATTTTCAAGCTACAAAGATGAAAGCTTTGGATGCTTTTG 480

QY 532 ATAGTGTCAAATTTAGTTTGGAGAGAGAGAACTATCAATGGCAATGCAAGTAT 591
DB 481 ATAGTGTCAAATTTAGTTTGGAGAGAGAGAACTATCAATGGCAATGCAAGTAT 540

QY 592 GGTGGCCAAAGTTTTCGAAATTAATTAATCACTGCGATGAGAGATGCAACGAGCCT 651
DB 541 GGTGGCCAAAGTTTTCGAAATTAATTAATCACTGCGATGAGAGATGCAACGAGCCT 600

QY 652 TTAACCTTTCGAAATTCGAAATTTGAAAGTGAATATATCAAGATTAATAATGCAAC 711
DB 601 TTAACCTTTCGAAATTCGAAATTTGAAAGTGAATATATCAAGATTAATAATGCAAC 660

QY 712 AAATTCATATCAAAATTTGAGTCAATGATATGTTAGCTTAAATTTGATGATCAATG 771
DB 661 AAATTCATATCAAAATTTGAGTCAATGATATGTTAGCTTAAATTTGATGATCAATG 720

QY 772 CTTGAGCAAGAAGCCCAATATAC 793
DB 721 CTTGAGCAAGAAGCCCAATATAC 742

```

```

RESULT 6
BM413343          749 bp  mRNA  linear  EST 22-JAN-2002
LOCUS            BM413343
DEFINITION      E18187670 tomato breaker fruit Lycopersicon esculentum cDNA clone
                  CLE63C11 5' end, mRNA sequence.

```

```

ACCESSION      EM413343
VERSION        BM413343.1
KEYWORDS       GI:18264973
SOURCE         EST
ORGANISM       Lycopersicon esculentum (tomato)
REFERENCE      Lycopersicon esculentum
AUTHORS        Bukaryota; Vitridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
                1 (bases 1 to 749)
                Alcalá, J., Vrebalov, J., White, R., Visions, T., Karamcheyeva, S. A.,
                Tsai, J., Bougri, O., Kirkness, E., Uteback, T., Van Aken, S.,
                Romling, C. M., Fraser, C. M., Martin, G. B., Tanksley, S. D. and
                Giovannoni, J.
                Generation of ESTs from tomato fruit tissue, breaker stage (2002)
                Unpublished (2002)
                Contact: CUGI

```

```

TITLE          JOURNAL
COMMENT         Clemson University Genomics Institute
                Clemson University
                100 Jordan Hall, Clemson, SC 29634, USA
                Email: http://www.genome.clemson.edu/orders/index.html
                This clone is available through the Clemson University Genomics
                Institute

```

```

FEATURES
source         Seq primer: T3;
                Location/Qualifiers

```

```

1..749
    /organism="Lycopersicon esculentum"
    /mol_type="mRNA"
    /cultivar="TA96"
    /db_xref="taxon:4081"
    /clone="CLE63C11"
    /issue_type="Pericarp"
    /dev_stage="Breaker"
    /lab_host="SOLR"
    /clone_lib="tomato breaker fruit"
    /note="Vector: pBluescriptSKcudap; Site 1: EcoRI;
    Site 2: XhoI; supplier: Boyce Thompson Institute;
    sequencing: The Institute for Genomic Research. Fruit
    were harvested at the breaker stage (first sign of
    lycopene accumulation on the blossom end of fruit). Fruit
    were cut in half and the seeds and locules were discarded
    prior to freezing the pericarp."

```

ORIGIN

```

Query Match      45.1%; Score 738; DB 12; Length 749;
Best Local Similarity 99.9%; Pred. No. 1.4e-113;
Matches 749; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 268 TTAAGTGAATTAATGATTTAGCTTTGAGCTTAAGGCTAATGAAATATGATATA 327
DB 1 TTAAGTGAATTAATGATTTAGCTTTGAGCTTAAGGCTAATGAAATATGATATA 60

QY 328 TTGCATTTAGCAAGATGAAATGAAAGCATGTCATCTGAACACCTGTTCAATTTGTG 387
DB 61 TTGCATTTAGCAAGATGAAATGAAAGCATGTCATCTGAACACCTGTTCAATTTGTG 120

QY 388 TTCTTAATAAACAAGATTTACTTCTCAACAATCACTTTTCAGGTCATGAGATCTT 447
DB 121 TTCTTAATAAACAAGATTTACTTCTCAACAATCACTTTTCAGGTCATGAGATCTT 180

QY 448 CTATTTCAATTAAGATTTTGGATCTTTCAGAACATCTAGTAATTTTCAAGTACAAAG 507
DB 181 CTATTTCAATTAAGATTTTGGATCTTTCAGAACATCTAGTAATTTTCAAGTACAAAG 240

QY 508 ATAGAAAGCTTTGATTTGCTTTTGAATGATCAAAATTTAGTTTGGAGAGAGAA 567
DB 241 ATAGAAAGCTTTGATTTGCTTTTGAATGATGATCAAAATTTAGTTTGGAGAGAGAA 300

QY 568 CTATCAATGGCAATGCAAGTATGCTGGCCAAAGTTCTTGCAAAATTAATTAATCACTGC 627
DB 301 CTATCAATGGCAATGCAAGTATGCTGGCCAAAGTTCTTGCAAAATTAATTAATCACTGC 360

QY 628 CATGCAAGGATGACCAACGAGCTTTAAGCTTTGGAATTTGCAAAATTTGAAGTGAATA 687

```

Db 361 CATCAGGAGATGACCAACGCGCTTAACTTCTGGAATTGCAAAATTGGAAGTGATA 420

Qy 688 ATCTAAAGAGTAAAGACACAAATTCATATCAATTTGACATNGCAGCTAATGTTG 747

Db 421 ATCTAAAGAGTAAAGACACAAATTCATATCAATTTGAGTCAGCTAATGTTG 480

Qy 748 TAGCTTCAAAATTTGATGATCAATGCTTCAGCAAGAGCCCAAAATATGATGAGTCATG 807

Db 481 TAGCTTCAAAATTTGATGATCAATGCTTCAGCAAGAGCCCAAAATATGATGAGTCATG 540

Qy 808 TATCAATATCTCAATATATTTCAATATCTGATATCTATTATTTGAAACAGCTATGATGTA 867

Db 541 TATCAATATCTCAATATATTTCAATATCTGATATCTATTATTTGAAACAGCTATGATGTA 600

Qy 868 TTTCAATTTGTTCTGGAATCTCAAAATGTCAGGCGCAAAATATTAATCTTGTCAGGTC 927

Db 601 TTTCAATTTGTTCTGGAATCTCAAAATGTCAGGCGCAAAATATTAATCTTGTCAGGTC 660

Qy 928 ATGGTATATGATTTGGAAGCTTAGAGCTGGAATTCAGAACTTATGTCATATGTTA 987

Db 661 ATGGTATATGATTTGGAAGCTTAGAGCTGGAATTCAGAACTTATGTCATATGTTA 719

Qy 988 CTGTAATGAAGCCAAATTTATCGTGCCG 1017

Db 720 CTGTAATGAAGCCAAATTTATCGTGCCG 749

RESULT 7

LOCUS BM413075 728 bp mRNA linear EST 22-JAN-2002

DEFINITION EST587402 tomato breaker fruit Lycopersicon esculentum cDNA clone

ACCESSION BM413075

VERSION BM413075.1 GI:18264705

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE Alcald, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A., Tsai, J., Bougri, O., Kirpness, E., Uteck, T., Van Aken, S., Roming, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

AUTHORS Giovannoni, J.

TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)

JOURNAL Unpublished (2002)

COMMENT Contact: CUGI

CONTACT CUGI

INSTITUTION Clemson University Genomics Institute

ADDRESS 100 Jordan Hall, Clemson, SC 29634, USA

EMAIL http://www.genome.clemson.edu/orders/index.html

NOTE This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

Location/Qualifiers

1..728

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA96"

/db_xref="taxon:4081"

/clone="CLEG62K21"

/issue_type="Pericarp"

/dev_stage="breaker"

/lab_host="SOLR"

/clone_id="tomato breaker fruit"

/note="Vector: pluscriptSKmCtadpt; Site 1: EcoRI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Query Match 44.4%; Score 726.4; DB 12; Length 728;

Best Local Similarity 99.9%; Pred. No. 1.2e-111;

Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 787 CAAATCTGATGAGTCCATGATATCAATATCTCAATATATTCATATCTATTA 846

Db 1 CAAATCTGATGAGTCCATGATATCAATATCTCAATATATTCATATCTATTA 60

Qy 847 TTGGAACAGGTGATGATTTGTTTCAATTTGTTTCTGATCTCAAAATGTCAGGCGACAA 906

Db 61 TTGGAACAGGTGATGATTTGTTTCAATTTGTTTCTGATCTCAAAATGTCAGGCGACAA 120

Qy 907 ATATTACTTGTCGTCAGGTCATGATATTAATTTGGAAGCTTGGATCTGGAATTCAG 966

Db 121 ATATTACTTGTCGTCAGGTCATGATATTAATTTGGAAGCTTGGATCTGGAATTCAG 180

Qy 967 AAGCTTATGTCCTAATGTTTACTGTAATGAAGCCAAATTTATCGTCCGAAAATGGAG 1026

Db 181 AAGCTTATGTCCTAATGTTTACTGTAATGAAGCCAAATTTATCGTCCGAAAATGGAG 240

Qy 1027 TTAGGATCAACACTTGGGAGGAGATCTGACAAAGCTAGCAACATCAATTTCTGATG 1086

Db 241 TTAGGATCAACACTTGGGAGGAGATCTGACAAAGCTAGCAACATCAATTTCTGATG 300

Qy 1087 TGGAAATGCAAGAGCTTAAGTATCCATTAATTATAGACCAAACTATGTGATCGAGTTG 1146

Db 301 TGGAAATGCAAGAGCTTAAGTATCCATTAATTATAGACCAAACTATGTGATCGAGTTG 360

Qy 1147 AACCATGATACCAACAGTTTTCAGCAGTTCAAGTGAATAATGTCGATGAGAAATATCA 1206

Db 361 AACCATGATACCAACAGTTTTCAGCAGTTCAAGTGAATAATGTCGATGAGAAATATCA 420

Qy 1207 AGGCGACAAGTGCACAAGGTGCGCTAAATTTGATGTCAGCAACAACCTTTCATG 1266

Db 421 AGGCGACAAGTGCACAAGGTGCGCTAAATTTGATGTCAGCAACAACCTTTCATG 480

Qy 1267 AAGGAATTTATATGAGATTAATTAATTAGGAGGAAAGTGAAGAACATCAGAGGCTA 1326

Db 481 AAGGAATTTATATGAGATTAATTAATTAGGAGGAAAGTGAAGAACATCAGAGGCTA 540

Qy 1327 CGTCAAAATATGTCATTTTAAACAATGTCGAACATGTTACACACACTGCACTTCACTAG 1386

Db 541 CGTCAAAATATGTCATTTTAAACAATGTCGAACATGTTACACACACTGCACTTCACTAG 600

Qy 1387 AAATTTAGAGGATGAAGCTTTTGTATATTTATTTATTTATTTATTTATTTATTTAT 1446

Db 601 AAATTTAGAGGATGAAGCTTTTGTATATTTATTTATTTATTTATTTATTTATTTATTTAT 660

Qy 1447 TATGACGATATGATATATCAATTAACAATCTATCTATGATGATTAATTTATTA 1506

Db 661 TATGACGATATGATATATCAATTAACAATCTATCTATGATGATTAATTTATTA 720

Qy 1507 TTAATATG 1514

Db 721 TTAATATG 728

RESULT 8

AM222257 729 bp mRNA linear EST 18-MAY-2001

LOCUS EST299068 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA

DEFINITION clone CLEW018, mRNA sequence.

ACCESSION AM222257

VERSION AM222257.1 GI:6533941

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE Alcald, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A., Tsai, J., Bougri, O., Kirpness, E., Uteck, T., Van Aken, S., Roming, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

AUTHORS Giovannoni, J.

TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)

JOURNAL Unpublished (2002)

COMMENT Contact: CUGI

CONTACT CUGI

INSTITUTION Clemson University Genomics Institute

ADDRESS 100 Jordan Hall, Clemson, SC 29634, USA

EMAIL http://www.genome.clemson.edu/orders/index.html

NOTE This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

Location/Qualifiers

1..728

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA96"

/db_xref="taxon:4081"

/clone="CLEG62K21"

/issue_type="Pericarp"

/dev_stage="breaker"

/lab_host="SOLR"

/clone_id="tomato breaker fruit"

/note="Vector: pluscriptSKmCtadpt; Site 1: EcoRI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."


```

Db      61  GATCAATCTTCAGCAAGAGCCCAATATCATGATGATCCATGATCAATATCATATA 120
Qy      824  TATTCAAATATCTGATATATATATATGGAACAGGTGATGATATTCATATGTTCTCG 883
Db      121  TATTCAAATATCTGATATATATATGGAACAGGTGATGATATTCATATGTTCTCG 180
Qy      884  ATTCCAAATATGAGCGCCACAATATATCTGTTGTCAGGTCACTGATATGATTTGG 943
Db      181  ATTCCAAATATGAGCGCCACAATATATCTGTTGTCAGGTCACTGATATGATTTGG 240
Qy      944  AAGCTTAGATCTGGAATTCAGAAAGCTTATGTCCTAATGTTACTGTAATGAAAGCAA 1003
Db      241  AAGCTTAGATCTGGAATTCAGAAAGCTTATGTCCTAATGTTACTGTAATGAAAGCAA 300
Qy      1004  AATATGCGTCCGCAAAATGAGATGATCAAGCTTGCAAGCGAGATCTGCAAGC 1063
Db      301  AATATGCGTCCGCAAAATGAGATGATCAAGCTTGCAAGCGAGATCTGCAAGC 360
Qy      1064  TAGCAACATCAAAATTTCTGAATGTAATGCAAGACCTTAAGTATCCCATATATAGA 1123
Db      361  TAGCAACATCAAAATTTCTGAATGTAATGCAAGACCTTAAGTATCCCATATATAGA 420
Qy      1124  CCAAAACTATTTGATGATGAGTTGAACCATGATACACAGTTTTCAGAGTTCAAGTAA 1183
Db      421  CCAAAACTATTTGATGATGAGTTGAACCATGATACACAGTTTTCAGAGTTCAAGTAA 480
Qy      1184  AAAATGCGTGTATAGAAATATCAAGGGGCAAGTGCAACAAAGGGGCAATAAATTTGA 1243
Db      481  AAAATGCGTGTATAGAAATATCAAGGGGCAAGTGCAACAAAGGGGCAATAAATTTGA 540
Qy      1244  TTGAGAGCAAAACTTTCCATGATGAGAAATATATAGAGAAATATTAATTTAGTAGGG 1303
Db      541  TTGAGAGCAAAACTTTCCATGATGAGAAATATATAGAGAAATATTAATTTAGTAGGG 600
Qy      1304  AAGTGAAGAAACCATCAGAGGCTACGTGCAAAATGTCATTTTAAACAATGCTGAACATGT 1363
Db      601  AAGTGAAGAAACCATCAGAGGCTACGTGCAAAATGTCATTTTAAACAATGCTGAACATGT 660
Qy      1364  TACACCACTGCACTGCACTGATAGAAATTTGAGAGATGAGAGCTCTTTTG 1412
Db      661  TACACCACTGCACTGCACTGATAGAAATTTGAGAGATGAGAGCTCTTTTG 709

RESULT 10
AM442253 696 bp mRNA linear EST 18-MAY-2001
LOCUS EST311649 tomato fruit red ripe, TMU Lycopersicon esculentum cDNA
DEFINITION AM442253
VERSION AM442253
KEYWORDS
SOURCE
ORGANISM Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eubacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 696)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upoun,J., Hansen,T., Craven,M.B., Bowman,C.L., Alm,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D., and
Giovannoni,J.
REFERENCE Generation of ESTs from tomato fruit tissue
COMMENT Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES
Location/Qualifiers
1..696
/organism="Lycopersicon esculentum"

```

```

/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEN22614"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TMU"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni. Fruit were tagged at the
breaker stage (first sign of lycopen accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

ORIGIN
Query Match 42.4%; Score 694.4; DB 10; Length 696;
Best Local Similarity 99.9%; Pred. No. 2.7e-106;
Matches 695; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      349  ATGAAGCATGTTCACTATGAAACACCTGTTCAATTTGGTTCCTAAAGAAAGATTTC 408
Db      1  ATGAAGCATGTTCACTATGAAACACCTGTTCAATTTGGTTCCTAAAGAAAGATTTC 60
Qy      409  TTCTCAAGCAATACCTTTTCAGGTCCATGACAGATCTTATTTCAATAAGATTTTG 468
Db      61  TTCTCAAGCAATACCTTTTCAGGTCCATGACAGATCTTATTTCAATAAGATTTTG 120
Qy      469  GATCCTTAGAAGCATAGTAAATTTGACATCAAAAGATAGAGGCTTTGATTCCT 528
Db      121  GATCCTTAGAAGCATAGTAAATTTGACATCAAAAGATAGAGGCTTTGATTCCT 180
Qy      529  TTGATAGGTTCAAAATTTAGTGTGGAGAGAGAGAACTATCAATGGCAATGGACAG 588
Db      181  TTGATAGGTTCAAAATTTAGTGTGGAGAGAGAGAACTATCAATGGCAATGGACAG 240
Qy      589  TATGATGGCCAAAGTTCTTGCAAAATTAATCACTCCATGCGAGGATGACACCAAG 648
Db      241  TATGATGGCCAAAGTTCTTGCAAAATTAATCACTCCATGCGAGGATGACACCAAG 300
Qy      649  CCTTACCTTCTGGAATTCGAAAAATTTGAAAGTGAAATATCTAAAGATGAAAAATGAC 708
Db      301  CCTTACCTTCTGGAATTCGAAAAATTTGAAAGTGAAATATCTAAAGATGAAAAATGAC 360
Qy      709  AACCAATTCATATCAATTTAGATGATGACATATATGTTGCTTCAAAATTTGATGATCA 768
Db      361  AACCAATTCATATCAATTTAGATGATGACATATATGTTGCTTCAAAATTTGATGATCA 420
Qy      769  ATGCTTCAGCAAGAGCCCAATCTCTGATGAGTGCATGATCAAAATCTCAATATATTC 828
Db      421  ATGCTTCAGCAAGAGCCCAATCTCTGATGAGTGCATGATCAAAATCTCAATATATTC 480
Qy      829  AAATATCTGATATATATATGAAACAGGTGATGATGATTTCAATGTTCTCGATCTC 888
Db      481  AAATATCTGATATATATATGAAACAGGTGATGATGATTTCAATGTTCTCGATCTC 540
Qy      889  AAAATGAGCGCCACAATATTTCTTGTGTCAGGTCACTGATATAGTATTTGAAAGCT 948
Db      541  AAAATGAGCGCCACAATATTTCTTGTGTCAGGTCACTGATATAGTATTTGAAAGCT 600
Qy      949  TAGATCTGGAATATTCGAAGCTTATGATGCTAATGTTACTGTAATGAAAGCAAAATTA 1008
Db      601  TAGATCTGGAATATTCGAAGCTTATGATGCTAATGTTACTGTAATGAAAGCAAAATTA 660
Qy      1009  TCGGTGCCGAAATGAGATTAGATCAAGACTTGGC 1044
Db      661  TCGGTGCCGAAATGAGATTAGATCAAGACTTGGC 696

RESULT 11
AM442335

```

LOCUS AM442335 687 bp mRNA linear EST 18-MAY-2001
 DEFINITION EST111731 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
 clone cLEN22L13 5', mRNA sequence.
 ACCESSION AM442335
 VERSION AM442335
 KEYWORDS EST.
 SOURCE EST.
 ORGANISM Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 687)
 Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
 Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
 Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
 Giovannoni,J.
 TITLE Generation of ESTs from tomato fruit tissue
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.
 FEATURES
 source
 1..687
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLEN22L13"
 /tissue_type="pericarp"
 /dev_stage="red ripe (7-20 days post-breaker)"
 /clone_lib="tomato fruit red ripe, TAMU"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; supplier: Giovannoni; fruit were tagged at the
 breaker stage (first sign of Lycopersicon accumulation on the
 blossom end of the fruit) and harvested 7 days
 post-breaker (fully red-ripe), 10 days post breaker, and
 20 days post-breaker (over-ripe). 20 day fruit which
 showed external or internal signs of pathogenesis were
 discarded. Fruit were cut in half and the seeds and
 locules were discarded prior to freezing the pericarp."

Query Match 41.8%; Score 683.8; DB 10; Length 687;
 Best Local Similarity 99.7%; Pred. NO. 1.6e-104;
 Matches 685; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 274 TGATTATGACTTAGCTTGGAGCTAAGGATGAGAAACATATGATTAATTCAT 333
 DB 1 TGATTATGACTTAGCTTGGAGCTAAGGATGAGAAACATATGATTAATTCAT 60

QY 334 TTGAGCAAGCATGGAATGAAGCATGTTCACTAGAACACCTGTTCAATTTGGTCCCA 393
 DB 61 TTGAGCAAGCATGGAATGAAGCATGTTCACTAGAACACCTGTTCAATTTGGTCCCA 120

QY 394 AAAACAAGATTTCTTCTCAAGCAATCACTTTTCAGGTCATGAGATCTTCAATT 453
 DB 121 AAAACAAGATTTCTTCTCAAGCAATCACTTTTCAGGTCATGAGATCTTCAATT 180

QY 454 CAGTAAGAATTTTGGATCCTTAGAAGCATCTAGTAAATTTGAGACTCAAGATAGAA 513
 DB 181 CAGTAAGAATTTTGGATCCTTAGAAGCATCTAGTAAATTTGAGACTCAAGATAGAA 240

QY 514 GCGTTTGGATGCTTTTGGATGCTTCAAAATTTAGTTGGAGAGAGAGAACTATCA 573
 DB 241 GCGTTTGGATGCTTTTGGATGCTTCAAAATTTAGTTGGAGAGAGAGAACTATCA 300

QY 574 ATGGCAATGGAAGAATAGTGGCAAGTCTTGCAAAATTAATTAATCACTGCAATGCA 633
 DB 301 ATGGCAATGGAAGAATAGTGGCAAGTCTTGCAAAATTAATTAATCACTGCAATGCA 360

QY 634 GGGATGCACCAACGGCCTTAACCTTGTGCAATTTGCAAAATTTGAAGAATATCTTA 693
 DB 361 GGGATGCACCAACGGCCTTAACCTTGTGCAATTTGCAAAATTTGAAGAATATCTTA 420

QY 694 AGAGTAAATGACACCAAAATTCATATGCAATTTGAGTCATGATTTGAGCTT 753
 DB 421 AGAGTAAATGACACCAAAATTCATATGCAATTTGAGTCATGATTTGAGCTT 480

QY 754 CAAATTTGATGATCAATGCTTCAGCAAAAGCCCAATTCATGAGTCATGATTCAA 813
 DB 481 CAAATTTGATGATCAATGCTTCAGCAAAAGCCCAATTCATGAGTCATGATTCAA 540

QY 814 ATACTCAATATATTCATATATCTGATATGATTTGGAACAGGTGATGATTTGCA 873
 DB 541 ATACTCAATATATTCATATATCTGATATGATTTGGAACAGGTGATGATTTGCA 600

QY 874 TTGTTCTGATCTCAAAATGTCAGCCCAAAATTTCTTGTCAGGTCAGTATGTA 933
 DB 601 TTGTTCTGATCTCAAAATGTCAGCCCAAAATTTCTTGTCAGGTCAGTATGTA 660

QY 934 TAAGTATGGAAGCTTAGATCTGGA 960
 DB 661 TAAGTATGGAAGCTTAGATCTGGA 687

RESULT 12
 BM536089 703 bp mRNA linear EST 20-FEB-2002
 LOCUS BM536089
 DEFINITION EST589111 tomato breaker fruit Lycopersicon esculentum cDNA clone
 CLE671174 5' end, mRNA sequence.
 ACCESSION BM536089
 VERSION BM536089
 KEYWORDS EST.
 SOURCE EST.
 ORGANISM Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 703)
 Alcala,J., Vrebalov,J., White,R., Vision,T., Karamyheva,S.A.,
 Tsai,J., Bougri,O., Kirkness,E., Utechtack,T., Van Aken,S.,
 Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
 Giovannoni,J.
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 JOURNAL Unpublished (2002)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3
 FEATURES
 source
 1..703
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLE671174"
 /tissue_type="pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /clone_lib="tomato breaker fruit"
 /note="Vector: pBluescriptSktdap; Site 1: EcoRI;
 Site 2: XhoI; supplier: Boyce Thompson Institute;
 sequencing: The Institute for Genomic Research. Fruit
 were harvested at the breaker stage (first sign of
 Lycopersicon accumulation on the blossom end of fruit). Fruit
 were cut in half and the seeds and locules were discarded
 prior to freezing the pericarp."

ORIGIN

/clone_11b="tomato breaker fruit"
/note="Vector: pBluescriptSKCUDapt; Site 1: Scori;
Site 2: XhoI; supplier: Boyce Thompson Institute; Fruit
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

ORIGIN

```
Query Match          40.2%; Score 657.2; DB 12; Length 670;  
Best Local Similarity 98.8%; Pred. No. 4,4e-100;  
Matches 662; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY      8  TTCAATAGACAAGTTAAACCATACATATACATATATCATGTATCCAAAGGA 67  
      1  TTCAATAGACAAGTTAAACCATACATATACATATATCATGTATCCAAAGGA 60  
      68  TAGATTCCTCTCTCATTTATTTTGGCTTCATCAATTTCACTGTAGAGCAATGT 127  
      61  TAGATTCCTCTCTCATTTATTTTGGCTTCATCAATTTCACTGTAGAGCAATGT 120  
      128  TATGATGACAATTTATCAACAAGTTATGATATATCTTGACAGAAATTTGCTCA 187  
      121  TATGATGACAATTTATCAACAAGTTATGATATATCTTGACAGAAATTTGCTCA 180  
      188  TGATTTTCAGCTTATCTTCTTATTTGAGCAAAAATATGAAAGCAACAATAATATGA 247  
      181  TGATTTTCAGCTTATCTTCTTATTTGAGCAAAAATATGAAAGCAACAATAATATGA 240  
      248  CAAGGTGATAAAATGGGATTAAAGTATTAATGTAATGTAAGCTTGAAGCTTAAGGGTGA 307  
      241  CAAGGTGATAAAATGGGATTAAAGTATTAATGTAATGTAAGCTTGAAGCTTAAGGGTGA 300  
      308  TGGAAAAACATATGATATATTTGATTTGACCAAGCATGGAATGAGCATGTCATCTAG 367  
      301  TGGAAAAACATATGATATATTTGATTTGACCAAGCATGGAATGAGCATGTCATCTAG 360  
      368  AACACCTGTCATTTGGTTCCTAAACCAAGAAATATCTTCTCAAGCAAAATCACCTT 427  
      361  AACACCTGTCATTTGGTTCCTAAACCAAGAAATATCTTCTCAAGCAAAATCACCTT 420  
      428  TTCAGGTCATGCAAGATCTTCTATTTCAAGAAAGATTTTGGATCTTAGAAGCATCTAG 487  
      421  TTCAGGTCATGCAAGATCTTCTATTTCAAGAAAGATTTTGGATCTTAGAAGCATCTAG 480  
      488  TAAATTTCAACTAAGAAAGATAGAGGCTTGGATTTGATTTGATATGTTCAAAATTT 547  
      481  TAGAATTTCAACTAAGAAAGATAGAGGCTTGGATTTGATTTGATATGTTCAAAATTT 540  
      548  AGTTGTTGAGAGAGAGAACTATCAATGCAATGCAATGCAATGATGAGGCAAGTTCTTG 607  
      541  AGTTGTTGAGAGAGAGAACTATCAATGCAATGCAATGCAATGATGAGGCAAGTTCTTG 600  
      608  CAAATTAATTAATCACTGCGCATGCGAGGATGACCAACGCGCTTAACTTCTGGAATTG 667  
      601  CAAATTAATTAATCACTGCGCATGCGAGGATGACCAACGCGCTTAACTTCTGGAATTG 660  
      668  CAAATTTTG 677  
      661  CAAATTTTG 670
```

Search completed: July 27, 2004, 02:46:01
Job time : 4445 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:04:52 ; Search time 28.379 Seconds
(without alignments)
831.356 Million cell updates/sec

Title: US-10-018-604-2

Perfect score: 2390
Sequence: 1 MWIQRNSILLIIIFASSIS.....VPHCTSLIESEDEALYNY 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
- 4: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
- 5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/prodata/2/iaa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	457	6	5447867-1 Patent No. 5447867
2	1113	46.6	433	4	US-09-051-239A-2 Sequence 2, Appl1
3	1103	46.2	433	3	US-08-941-532-6 Sequence 6, Appl1
4	787	32.9	514	3	US-08-467-023-134 Sequence 134, App
5	365	15.3	70	6	5453566-2 Patent No. 5453566
6	319	13.3	127	3	US-08-467-023-189 Sequence 189, App
7	299	12.5	451	4	US-09-107-532A-6652 Sequence 6652, App
8	298.5	12.5	452	1	US-08-290-878A-5 Sequence 5, Appl1
9	298.5	12.5	452	2	US-08-780-869-5 Sequence 2, Appl1
10	293.5	12.3	362	4	US-09-787-583-2 Sequence 2, Appl1
11	257.5	10.8	360	3	US-09-542-767A-1 Sequence 1, Appl1
12	241	10.1	440	1	US-08-061-062A-6 Sequence 6, Appl1
13	241	10.1	440	1	US-08-061-062A-8 Sequence 8, Appl1
14	241	10.1	440	3	US-08-536-150-6 Sequence 6, Appl1
15	241	10.1	440	3	US-08-536-150-8 Sequence 8, Appl1
16	226.5	9.5	127	3	US-08-467-023-188 Sequence 188, App
17	212.5	8.9	442	4	US-09-107-532A-6254 Sequence 6254, App
18	205.5	8.6	415	4	US-09-198-956-6 Sequence 6, Appl1
19	205.5	8.6	415	4	US-09-670-141-6 Sequence 187, App
20	188.5	7.9	128	3	US-08-467-023-187 Sequence 187, App
21	187	7.8	49	3	US-08-941-532-8 Sequence 8, Appl1
22	129.5	5.4	1005	4	US-09-206-942-41 Sequence 41, Appl1
23	129.5	5.4	1011	4	US-09-206-942-39 Sequence 39, Appl1
24	124.5	5.2	901	4	US-09-134-001C-5361 Sequence 5351, App
25	120.5	5.0	1004	4	US-09-206-942-57 Sequence 57, Appl1
26	120.5	5.0	1010	4	US-08-942-55 Sequence 55, Appl1
27	120	5.0	1535	3	US-08-755-587-185 Sequence 185, App

28	118.5	5.0	1073	4	US-09-206-942-49	Sequence 49, Appl1
29	118.5	5.0	1079	4	US-09-206-942-47	Sequence 47, Appl1
30	117	4.9	378	4	US-09-134-001C-4692	Sequence 4692, App
31	116.5	4.9	977	4	US-09-206-942-53	Sequence 53, Appl1
32	116.5	4.9	983	4	US-09-206-942-51	Sequence 51, Appl1
33	115	4.8	942	4	US-09-489-039A-13982	Sequence 13982, App
34	114	4.8	1036	4	US-09-206-942-73	Sequence 73, Appl1
35	114	4.8	1477	4	US-09-206-942-71	Sequence 71, Appl1
36	113.5	4.7	969	4	US-09-206-942-32	Sequence 32, Appl1
37	113.5	4.7	975	4	US-09-206-942-30	Sequence 30, Appl1
38	112	4.7	1612	1	US-08-169-927-2	Sequence 2, Appl1
39	110	4.6	975	4	US-09-328-352-4764	Sequence 4764, App
40	110	4.6	1220	4	US-09-206-942-28	Sequence 28, Appl1
41	110	4.6	1226	4	US-09-206-942-26	Sequence 26, Appl1
42	109.5	4.6	1338	2	US-08-728-470-9	Sequence 9, Appl1
43	109.5	4.6	1338	2	US-08-719-641-9	Sequence 9, Appl1
44	109.5	4.6	1529	2	US-08-728-470-10	Sequence 10, Appl1
45	109.5	4.6	1529	3	US-08-719-641-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1	5447867-1	Patent No. 5447867
APPLICANT: BRIDGES, IAN; SCHUCH, WOLFGANG; GRIERSON, DONALD		
TIME OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN		
RESTRICTION GENE SEGMENTS		
NUMBER OF SEQUENCES: 4		
CURRENT APPLICATION DATA:		
APPLICATION NUMBER: US/08/24,866		
FILING DATE: 26-FEB-1993		
PRIOR APPLICATION DATA:		
APPLICATION NUMBER: 720,629		
FILING DATE: 25-JUN-1991		
APPLICATION NUMBER: 419,779		
FILING DATE: 29-SEP-1989		
APPLICATION NUMBER: 119,614		
FILING DATE: 12-NOV-1987		
SEQ ID NO:1:		
LENGTH: 457		
5447867-1		
Query Match	100.0%;	Score 2390; DB 6; Length 457;
Best Local Similarity	100.0%;	Pred. No. 5.8e-208;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MWIRNSILLIIIFASSISITCRSNVIDNLFKQVYDNIIEFPFADPQAYLSYLSKNIE 60
DB	1	MWIRNSILLIIIFASSISITCRSNVIDNLFKQVYDNIIEFPFADPQAYLSYLSKNIE 60
QY	61	NNNNIDKVNKGKIVNVISFGAKGDKTYDNIAPFQAWMEACSSRPVQFVPPKNNYL 120
DB	61	NNNNIDKVNKGKIVNVISFGAKGDKTYDNIAPFQAWMEACSSRPVQFVPPKNNYL 120
QY	121	LKQITSPGCRSSISYKIFGSLSEASSKISDYKDRRLMIAFDSYQNLVWGGGTTNGQV 180
DB	121	LKQITSPGCRSSISYKIFGSLSEASSKISDYKDRRLMIAFDSYQNLVWGGGTTNGQV 180
QY	181	MWSSCKINKSLPCRDAPALATFWNCKNLKNNLKSKNAQOIHKEKSCNNVVASNNMIN 240
DB	181	MWSSCKINKSLPCRDAPALATFWNCKNLKNNLKSKNAQOIHKEKSCNNVVASNNMIN 240
QY	241	ASAKSPNTGVHVSNNQYIOISDTITGTGDDCSIVSGSNVQANITTCPPGHISIGSL 300
DB	241	ASAKSPNTGVHVSNNQYIOISDTITGTGDDCSIVSGSNVQANITTCPPGHISIGSL 300
QY	301	GGNSSEAYVSNVTVAEAKTIGANGVRIKTMQSGGQASNIKFLNTEMQDVKPIIIDQN 360
DB	301	GGNSSEAYVSNVTVAEAKTIGANGVRIKTMQSGGQASNIKFLNTEMQDVKPIIIDQN 360
QY	361	YCDRVEPCIQGSAVOVKNVVENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGESG 420

Db 361 YCDREVEPCIOQFSAVQVNVVYENIKGTSATKVAIKFDCSNPCEGIMENINLVGSG 420
QY 421 KPSBATCKNVHFNNAEHVTPEHCTSLSEI SEDBALVNY 457
Db 421 KPSBATCKNVHFNNAEHVTPEHCTSLSEI SEDBALVNY 457

RESULT 2

US-09-051-239A-2
Sequence 2, Application US/09051239A
Patent No. 6420628
GENERAL INFORMATION:
APPLICANT: ULVSKOV, Peter
APPLICANT: CHILD, Robin
APPLICANT: VAN ONCKELIN, Henri
APPLICANT: PRINSEN, Els
APPLICANT: BORKHARDT, Bernard
APPLICANT: SANDER, Ilija
APPLICANT: PETERSEN, Morten
APPLICANT: BUNDGARD POULSEN, Gert
APPLICANT: BOTTERMAN, Johan
TITLE OF INVENTION: Seed Shattering
FILE REFERENCE: 2121-0138P
CURRENT APPLICATION NUMBER: US/09/051,239A
CURRENT FILING DATE: 1998-09-28
PRIOR APPLICATION NUMBER: PCT/EP96/04313
PRIOR FILING DATE: 1996-10-04
PRIOR APPLICATION NUMBER: EP 95 402241.4
PRIOR FILING DATE: 1995-10-06
PRIOR APPLICATION NUMBER: EP 95 203328.0
PRIOR FILING DATE: 1995-12-08
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 433
TYPE: PRT
ORGANISM: Brassica napus
FEATURE:
OTHER INFORMATION: Strain cv. Topaz.
US-09-051-239A-2

Query Match 46.6%; Score 1113; DB 4; Length 433;
Best Local Similarity 49.9%; Pred. No. 2,2e-92;
Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

QY 7 SILLIIIFASISITCSNSVIDNLFKOVYNIIEQFAHDFQAYLSLKNISNNND 66
Db 9 AIFLCVILMLACCOALSSNV-DDG-----YGHEDG---SFESDSLTKLNDD 51
QY 67 KV-----DKNGIKYINVLSPFGAKDGKTYDNIAPQAMNEACSSRTPVQFVVPKKN 118
Db 52 DVLLTKSDRPTSSSTYSVSNFPAKDGKTDYQAFKAMKAKCSTNGVTFILPKGKT 111
QY 119 YLLKQITFSGPCRSSISVKTIFGSLBASISDY-KDRRLWIAPDSVONLVVGG--GTIN 175
Db 112 YLLKSIIFRFGPCRSIRSFQILGTLSASTKRSYSDNKKHMLILDVNNLSIDGSAGIVD 171
QY 176 GNGQVWPPSSCKINKSLPCRPAPALTFMCKNLKVNLSKNAQOIHIFESCTNYVAS 235
Db 172 GNGNIMWQNSCKIDKSPCTAPALTLNMLNKNLVRNAQOIOISTEKNNVGVK 231
QY 236 NLMINASAKSPNTDGVHVSNTQIYISDTIIGTDDCISIVSGSONVOATYTCGPBGHI 295
Db 232 NVKTLAEDGSNTGTGIRIVATKIRISNSDGTGDDCISIDGSQNVQINDLTGPGHGI 291
QY 296 SIGSLGSGSEAYVSYNTVNEAKIIGAENGRIKTWQSGSQASNIKPLANVEMQVXPI 355
Db 292 SIGSLGSDNSKAYVSGIDVDGATLSETDNGVRIKTYOQSGSTAKNIKFNQIRMDNVKPI 351
QY 356 IIDQNYCDREVEPCIOQFSAVQVNVVYENIKGTSATKVAIKFDCSNPCEGIMENINL 415
Db 352 IIDQNYCDK-DKCEQOESAVQVNVVYONIKGTSATVAIIMNCSVXVPCGIVLENNI 410

QY 416 VGESGKPSBATCKNVHFNNAEHVTPEH 442
Db 411 KG-----GKASCENNVYDKGTVSEPKC 432

RESULT 3

US-08-941-532-6
Sequence 6, Application US/08941532
Patent No. 6096946
GENERAL INFORMATION:
APPLICANT: ROBERTS, Jeremy Alan
APPLICANT: COUPE, Simon Allan
APPLICANT: JENKINS, Elizabeth Sarah
TITLE OF INVENTION: CONTROL OF POD DEHISCENCE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941,532
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00757
FILING DATE: 29-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9506684.1
FILING DATE: 31-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Besmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0623.0580001/RWE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-941-532-6

Query Match 46.2%; Score 1103; DB 3; Length 433;
Best Local Similarity 49.8%; Pred. No. 1,7e-91;
Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

QY 7 SILLIIIFASISITCSNSVID-----DNLFKOVYDNLIEQFAHDFQAYLSLKNIS 61
Db 9 AIFLCVILMLACCOALSSNVDDYGHEDGSFET--DSLTKLNDDVLLTKSDRPTTES 66
QY 62 NNNIDKYDKNGIKYINVLSPFGAKDGKTYDNIAPQAMNEACSSRTPVQFVVPKKNYL 121
Db 67 S-----TVSVSNFPAKDGKTDYQAFKAMKAKCSTNGVTFILPKGTYLL 114
QY 122 KQITFSGPCRSSISVKTIFGSLBASISDY-KDRRLWIAPDSVONLVVGG--GTINAG 178
Db 115 KSIRFRGPCRSIRSFQILGTLSASTKRSYSDNKKHMLILDVNNLSIDGSAGIVDNG 174
QY 179 QVWPPSSCKINKSLPCRPAPALTFMCKNLKVNLSKNAQOIHIFESCTNYVASLML 238
Db 175 KIMWQNSCKIDKSPCTAPALTLNMLNKNLVRNAQOIOISTEKNVSDVKNV 234

```

QY      INASKSNTGCVHVSNNQYVQISPTTIGTGDDCISYSGSQNVAQTITCGPHGISG 298
        |||:||||::||::||::||::||::||::||::||::||::||::||::||
Db      ITAPEDSNIDGIHVATKNIRISNSDGTGDCCISTIEDSQNVQINDLTCTGPCHGISG 294
QY      SLGSGNSEAVYSNVTWNEAKTIGAENGVRIKTWGSGSGOASINIFLWEMQDVXPIIID 358
        |||:||||::||::||::||::||::||::||::||::||::||::||::||
Db      SLGGDNSSAYYSGINVDATLSFTPDNGVRITTYOGSGSTAGNIIFQNRIMNVNXPFIID 354
QY      QNYCDRVEPCTIQFSAAVOVKNVYENIKGTSATKVAIKFDGSTNPFCEGIMENINLWGE 418
        |||:||||::||::||::||::||::||::||::||::||::||::||::||
Db      QNYCDK-DKCEQSAAVVNNVVYRNIIQTSATDVAIMFNCVAVPCQGITLENVNITKG- 412
QY      SGKPESEATCKVNHENNAEHTPHC 442
        :||:||||:~::~||:~:
Db      413 ----GKASCXKNVNWKDGTATSPKC 432

RESULT 4
US-08-467-023-134
; Sequence 134, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griffith, Irwin J.;
; APPLICANT: Pollock, Joanne;
; APPLICANT: Bond, Julian P.;
; APPLICANT: Garman, Richard D.;
; APPLICANT: Kuo, Mei-Chang;
; APPLICANT: Yeung, Siu-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; TITLE OF INVENTION: Japanese Cedar Pollen
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 514 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-467-023-134

Query Match          32.9%; Score 787; DB 3; Length 514;
Best Local Similarity 40.5%; Pred. No. 9,1e-63;
Matches 168; Conservative 73; Mismatches 160; Indels 14; Gaps 7;
45 AHDFAYLV---SYLSKNTESNNNDIKVD---KNGIKIVTNVLSFGAKDGKGYNDYNIAFEQA 98

```

```

Db      22 AEDGSAQIMLDVYVEKILRNRSLRKVHEHSHRDAINIFNVEIKYCAVGDGKHDCTEARSTA 181
      99 WNEACSSRTVPQFVVPKKNVYLKQITFSGPCSSISVKIFGSLIASSKISDYKDRBLMT 158
      82 WQACCKNPS-AMLVPESSKFFVNVNLFNFGPCQPHFFFKYDGTILAAVQNPASMKNNRIMWL 140
Qy      159 AFDSVQNLVVGGGGTINGNGQVWPFSSCK-INKSLPC--RDAPALAFMNCNLKVNMLX 215
Db      141 QFALTLFTLMGKGVLDIGQCKQWMAQQCKWNNGREICNDRDRPTAIFDFSTGLIIQGLK 200
Qy      216 SKNAQQHIFESCTNVVNASNMLMTNASKSPNTDGVHVSNTQYIQISDITLITGDDDCISI 275
Db      201 LMSNPEHLVFNGBEGYKIIIGISTTAPRDSPTVDGIDIFASKNPHLCKNTTIGTDDCVAI 260
Qy      276 VSGSQVQANVNTITGCPHGHSIGSLGSGNSEAVYNSVTVNEAKLIGAEVRIKWTQGGGS 335
Db      261 GTGSSNVIEDLIGCPHGHSISGLGNGENRAEVSYHVHNGAKRIDQNGRLRITWQGGGS 320
Qy      336 GQASNIKFLVEMQDVKKYPIIIDQNYCDRVEPCIQGFSAVOVKNVYENIKGISATKVAI 395
Db      321 GMASHIIEYENWEMINSNPILINQFYCTSSASACONRSAYQIDVTVYKINRISATPAAI 380
Qy      396 KFDSTNPPCGGIMENINLVGSGKSPATCKKVHFNN--AEVHTPCTSLETS 448
Db      381 QLKCSDSNPKDILKLSIDSLKLTSGK--IASCLNDNANGVFSGHVITPACKNLSFS 433

RESULT 5
5453566-2
/ Patent No. 5453566
/ APPLICANT: SHEWMAKER, CHRISTINE K.; KRIDL, JEAN C.; HATT,
/ WILLIAM R.; KNAUF, VIC
/ TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
/ IN PLANT/CELLS
/ NUMBER OF SEQUENCES: 2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/750,505
/ FILING DATE: 27-AUG-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 240,408
/ FILING DATE: 30-AUG-1988
/ APPLICATION NUMBER: 920,574
/ FILING DATE: 17-OCT-1986
/ APPLICATION NUMBER: 845,676
/ FILING DATE: 28-MAR-1986
/ SEQ ID NO:2
/ LENGTH: 70
5453566-2

Query Match      15.3%; Score 365; DB 6; Length 70;
Best Local Similarity 100.0%; Pred. No. 7.7e-26;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      72 GIXVINVLSFGAKDGGKTYNDINIAFEQAMNEACSSRTVPQFVVPKKNVYLKQITFSGPCR 131
Db      1 GIXVINVLSFGAKDGGKTYNDINIAFEQAMNEACSSRTVPQFVVPKKNVYLKQITFSGPCR 60

Qy      132 SSISVKIFGS 141
Db      61 SSISVKIFGS 70

RESULT 6
US-08-467-023-189
/ Sequence 189, Application US/08467023
/ Patent No. 6090386
/ GENERAL INFORMATION:
/ APPLICANT: Griffith, Irwin J.;
/ APPLICANT: Pollock, Joanne;
/ APPLICANT: Bond, Julian F.;
/ APPLICANT: Garman, Richard D;
/ APPLICANT: Kuo, Mei-Chang;

```

APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allegenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMT-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 189:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-467-023-189

Query Match 13.3%; Score 319; DB 3; Length 127;
Best Local Similarity 56.0%; Pred. No. 2.7e-21;
Matches 65; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

QY 293 HGISIGLSGNSPAAVSNVTNEAKTIIGAENGVRITKMGSGQASNIKFLNTEMQDVK 352
DB 8 HEFISIGLSGNSPAAVSNVTNEAKTIIGAENGVRITKMGSGQASNIKFLNTEMQDVK 67

QY 353 YPIIIDNVCYDRVPCIQQFSANQVKNVVENINGTSATKAKAFDDSTNFPCEGI 408
DB 68 NPIIINQFYCTASACQNRSAVQIDVTYKNINGTSATAAIAIQLKCSDSYPCDOI 123

RESULT 7
US-09-107-532A-6652
Sequence 6652, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 6652:
SEQUENCE CHARACTERISTICS:
LENGTH: 451 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...451
SEQUENCE DESCRIPTION: SEQ ID NO: 6652:
US-09-107-532A-6652

Query Match 12.5%; Score 299; DB 4; Length 451;
Best Local Similarity 27.2%; Pred. No. 1.1e-18;
Matches 99; Conservative 65; Mismatches 132; Indels 68; Gaps 15;

QY 73 IKVINIVSFGAKGDKTYNDIAFEQAMNEACSSRTPVQFVYKKKNYLLKQITFSGPCRS 132
DB 5 IMYDILTFGASID--ELNTEALIQADIDAAISDGGT--VVVPAGE-----FLTGALFL 54

QY 133 SISYKIFGSLSEASSKIS-DYKD-----RLW---IARDSYQNLVVGSGGTG 176
DB 55 KSNVEHLISGAVALKFDDEPDYVHVRMEGVARKYVASCTIYQNVENISVTGGLTD 114

QY 177 NGQVW-----PSCINKSLPCRDAPFALTFWNCKRLKYNLKSXAQOIHKFESCT 230
DB 115 NGKKWHTFFRNEPD---NLAYP---RPKLSFHNCHRITVXDIKLIQSPMTINPIICS 167

QY 231 NVVASNIMINASAPPTDGVHVSNTQYIOISDRIIGTDDCISIVGS-----QN 281
DB 168 NATFDNLILNPADSPVTDGIDPESCKVRLSNCHIDVGDCLAKGTEETTERIACEN 227

QY 282 VQATNITCGPGH-GISIGLSGNSPAAVSNVTNEAKTIIGAENGVRITKMGSGQASN 340
DB 228 ITITNCTMVHGGVGLVSEMSGS---INRITISNCIFQETDGRILKSRGGGIYED 283

QY 341 IKPLNEMQDVKYIITIDQNY-CD--RVEFCIQ-----PSAVQKVVYENI 385
DB 284 IRVSNIVDVNMCPIILNLYFCGPRGKPEYVWKKAYPIDERTPARIRHFNSITARNV 343

QY 386 KGTG 389
DB 344 HASA 347

RESULT 8
US-08-290-978A-5
Sequence 5, Application US/08290978A
Patent No. 5624834
GENERAL INFORMATION:
APPLICANT: KUSTERS-VAN SOMEREN, MARGO A.
APPLICANT: MULLER, YVONNE

APPLICANT: KESTER, HERMANUS C.M.
APPLICANT: VISSER, JACOB
APPLICANT: VAN COYEN, ALBERT J.J.
APPLICANT: ROLIN, CLAUD
TITLE OF INVENTION: CLONING AND EXPRESSION OF THE
TITLE OF INVENTION: EXO-POLYGALACTURONASE GENE FROM ASPERGILLUS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,978A
FILING DATE: 17-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 4615-0044.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-978A-5

Query Match 12.5%; Score 298.5; DB 1; Length 452;
Best Local Similarity 25.1%; Pred. No. 1.3e-18;
Matches 107; Conservative 73; Mismatches 166; Indels 81; Gaps 16;

74 KVINLVSFGAKGDKTYDNIAPFOANNEACSSRTPOVFPKKNKYL--KQITFSGPC 130
48 KTCVASHGEGTDDSYILSALNQ-----CNHGKGVF--DEKKEYILGALNMTF---- 96
131 RSSISVKIFGSL-----EASSKISDYKRRRLMTAFDSVONLVVGGGTTNGNQVW 181
97 LKNIDLEVLTITLFTNDTDYWCANSFKQGFQNAATTFQLOG-EDVNMVGGGTTNGNQVW 155

182 WPSCKINKSLPRDAPTLALTFWNCKNKKVNNKSKNAQOIHKFESCTNVVSNMUNA 241
156 YDIYADDDLTL---PPIIMGIIIGLNGGTIGPKLRISPOYHFFVANSNVLPFGIDISG 211
242 SAKS-----PNTDGVHVSNTQYIQISDTIIIGTGDICISIVSGSQNVQATNITCGPGHGISI 297
212 YKSDNDKAKNTDMDYRSMNIVIONSVINNGDDCVSFKRNSNIIIVQNHLHGNSHGISV 271
298 GSISGSGNSE-----AYVSNVTVNEAKIIGAENG-----RIKTWQ----- 332
272 GSISGQKXDEVDIAVENYVYVNIISMFNASVCINFNHIIIDFLTLTWLODMARIKIVWPGTSALS 341
333 -----GGSGQANIKPLINWEMODVKKPILIIDQNYCR--VEPCIQGSAVQVKNVYENI 385
332 ADIQGGGSGSVKNTIYDVALIDNVDALEITOCYGGKNTLLCNRPSSSLTISDVHAKNF 391
386 KGTSA-----TKVAIKFDCSTNPFCEGIIIMENINLVGESGKPSKSEATCKNVHFNNAEH 437
392 RGTTSSESDPYVGTIV-----GSSPDTCSDIYSNINIVTSPDG--TNDFCVDNV---DESL 442
438 VTPHGTS 444
:::|::|:

Db 443 LSVNCTA 449

RESULT 9
US-08-780-869-5
Sequence 5, Application US/08780869
Patent No. 5830737
GENERAL INFORMATION:
APPLICANT: KUSTERS-VAN SOMEREN, MARGO A.
APPLICANT: MULLER, YVONNE
APPLICANT: KESTER, HERMANUS C.M.
APPLICANT: VISSER, JACOB
APPLICANT: VAN COYEN, ALBERT J.J.
APPLICANT: ROLIN, CLAUD
TITLE OF INVENTION: CLONING AND EXPRESSION OF THE
TITLE OF INVENTION: EXO-POLYGALACTURONASE GENE FROM ASPERGILLUS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780,869
FILING DATE: 24-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/290,978
FILING DATE: 17-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 4615-0044.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-780-869-5

Query Match 12.5%; Score 298.5; DB 2; Length 452;
Best Local Similarity 25.1%; Pred. No. 1.3e-18;
Matches 107; Conservative 73; Mismatches 166; Indels 81; Gaps 16;

74 KVINLVSFGAKGDKTYDNIAPFOANNEACSSRTPOVFPKKNKYL--KQITFSGPC 130
48 KTCVASHGEGTDDSYILSALNQ-----CNHGKGVF--DEKKEYILGALNMTF---- 96
131 RSSISVKIFGSL-----EASSKISDYKRRRLMTAFDSVONLVVGGGTTNGNQVW 181
97 LKNIDLEVLTITLFTNDTDYWCANSFKQGFQNAATTFQLOG-EDVNMVGGGTTNGNQVW 155

182 WPSCKINKSLPRDAPTLALTFWNCKNKKVNNKSKNAQOIHKFESCTNVVSNMUNA 241
156 YDIYADDDLTL---PPIIMGIIIGLNGGTIGPKLRISPOYHFFVANSNVLPFGIDISG 211
242 SAKS-----PNTDGVHVSNTQYIQISDTIIIGTGDICISIVSGSQNVQATNITCGPGHGISI 297
212 YKSDNDKAKNTDMDYRSMNIVIONSVINNGDDCVSFKRNSNIIIVQNHLHGNSHGISV 271
298 GSISGSGNSE-----AYVSNVTVNEAKIIGAENG-----RIKTWQ----- 332


```

1      FILING DATE:  14 MAY 1993
2      CLASSIFICATION:  435
3      ATTORNEY/AGENT INFORMATION:
4      NAME:  KOKULIS, PAUL N.
5      REGISTRATION NUMBER:  16773
6      REFERENCE/DOCKET NUMBER:  202390/R 7262 (V)
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE:  (202) 861-3000
9      TELEFAX:  (202) 823-0944
10     TELEX:  6714627 CUSH
11
12     INFORMATION FOR SEQ ID NO:  6:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH:  440 amino acids
15     TYPE:  amino acid
16     TOPOLOGY:  linear
17
18     MOLECULE TYPE:  protein
19
20     US-08-061-062A-6

```

Query Match	10.1%;	Score 241;	DB 1;	Length 440;
Best Local Similarity	24.0%;	Pred. No. 2e-13;		
Matches	88;	Conservative	61;	Mismatches 186;
			Indels	32;
			Gaps	11

```

QY 70 KNGIIVIVLSTFGAGDGGKTYDNIAFEQAMNACSSRTPEVQVFNPKNKOYILKQITEFSGP 129
Db 33 KGATKTCNLTSGVAYADNSTVDGPATLSAW-AAKCKSGGLV--YIPSGNALNLTWVLTITGG 89
QY 130 CRSSISVK--IFGSLHASSKISDYKDRMLIAFDYQNTLVGGGGGTINGNGQYMWBPSSCK 187
Db 90 SATALIQDGLIIVRTGTASGNMLAVDTTDPLEFSSYISK-----GAVQGFQYVYHAEGLY 143
QY 188 INKSLPCRPAPALTFNMCNKLKANLKKRAQOHKEEBCSTNVVASNLMTASAKSPN 247
Db 144 GARILRLTDV-----THFSVHVLIVDPAPEFHTMDTSCDGHVYNNALRGNEG-G 193
QY 248 TDGVHVSNTQYQISPTIIGTDDPCISIVSGSNQVATNITGPRGHGISIGLSGNSSEA 307
Db 194 LDGIDIVWGSN-IWVHDVEYTNKDECVTASPAVNIIVESIYCNWSGGCAAGSLG--ADT 249
QY 308 YVSNVTVAEAKLIGANGVRIKTWQGSQASNIKEPLANEMODVKPIIIDQYVCDRVEP 367
Db 250 DVTLDVYNNVWTSSNQWMIKS-NQGSSTVSNVLEENIIGHONALSLDIDGYWSSMTAV 308
QY 368 CIQGFSAVQKNVYENIKGT--SATKYAIFKFGSTNPPCEGIIMENINLVGSEKPSF 424
Db 309 A---GDGYQLNNITVKNMKGTFRANGATRPPIRVVCSDTAPCTDLTIEDIAIWTSGSSSEL 365
QY 425 ATCKNVH 431
Db 366 YLCRSAY 372

```

RESULT 13
US-08-061-062A-8

Sequence 8, Application US/06010624
Patent No. 5550045
GENERAL INFORMATION:
APPLICANT: MUSTERS, WOUTER
APPLICANT: STAM, HEIN
APPLICANT: SUYKERBOUY, MARIA E.
APPLICANT: VISSER, JACOB
APPLICANT: VERBAKEI, Johannes M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
TITLE OF INVENTION: ENCODING A RIPENING FORM OF A POLYPEPTIDE HAVING
TITLE OF INVENTION: RHANOGALACTURONASE ACTIVITY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk

```

1      COMPUTER:  IBM PC compatible
2      OPERATING SYSTEM:  PC-DOS/MS-DOS
3      SOFTWARE:  PatentIn Release #1.0, Version 1
4      CURRENT APPLICATION DATA:
5      APPLICATION NUMBER:  US/08/061,062A
6      FILING DATE:  14 MAY 1993
7      CLASSIFICATION:  435
8      ATTORNEY/AGENT INFORMATION:
9      NAME:  KOKULIS, PAUL N.
10     REGISTRATION NUMBER:  16773
11     REFERENCE/DOCKET NUMBER:  202390/R 7262 (V)
12     TELECOMMUNICATION INFORMATION:
13     TELEPHONE:  (202) 861-3000
14     TELEFAX:  (202) 822-0944
15     TELEX:  6714627 CUSH
16     INFORMATION FOR SEQ ID NO:  8:
17     SEQUENCE CHARACTERISTICS:
18     LENGTH:  440 amino acids
19     TYPE:  amino acid
20     TOPOLOGY:  linear
21     MOLECULE TYPE:  protein
22     US-08-061-062A-8

```

Query March	10.1%	Score 241;	DB 1;	Length 440;
Best Local Similarity	24.0%	Pred. No. 2e-13;		
Matches 88;	Conservative 61;	Mismatches 186;	Indels 32;	Gaps 11.

```

QY      70 KNGIKYINVLISGAKGDGKTYDNTIAFEQAWNEACSSBTPVQFYVVKPKKNYLLKQIFPSG 129
      33 KGATKTCNLLISGVAANDSTVDGPILTSAM--AKKSGGLV--YIPSGYALNTWVLLTGG 89
QY      130 CRESSIEVK--IRGSLI EASSKSIDYKDRFLMIAFDVSQNLVYGGGGLTNGNQOYVWPSSCK 187
      90 SATALIDBQIILYRTGTAAGNMIAVYTDLPDFELFESTSK-----GAYQGGYVYHAGCY 143
QY      188 INKSLPECDAPALTFENCKRLKVNNEKSKQAQOIHIFKESCTNVVANSNIMINSAKSPN 247
      144 GARILRLTDV-----TFHSYEDVLVADPAFHFMTDCTSGEYVYNAINGNNG-G 193
QY      248 TGCYVHSNTQYQISDITLIIGDPCDISVSSSQVQAMNTTCGGGHSISGSLGSGNSEA 307
      194 LQGDIDWGSN-LWVHDEVVTKDECVYKSPANNILVESIYICMWSGGCAGMSIGS--ADT 249
QY      308 YVSNVTVNENAKIIGAGENVRIKTQWGGSGQASNIKFLVNEQVDKYPIIIDQVYCDREVP 367
      250 DVTYDLYRRVRYVWSSNQWYMIKS--NGSGGTYSNVLLFNIGHGAYSLDLDIDGWSSMTAV 308
QY      368 CIOQFSAVQKVVYENIKGT--SAIKVAIKFDQSTNPFCEGIMENINLVGSGKPS 424
      309 A--GDGVQNNIITVKWKKGTGANATRPPIRVVCSDPATCTDLIEDIAIMTESGSSSL 365
QY      425 ATCKNVH 431
      366 YLCRSAY 372
Db

```

RESULT 14
US-08-536-150-6

Sequence 8, Application US/06356130
Patent No. 6013489
GENERAL INFORMATION:
APPLICANT: MUSTERS, MOUTER
APPLICANT: STAM, HEIN
APPLICANT: SUYKERBUIJK, MARIA E.
APPLICANT: VISSER, JACOB
APPLICANT: VERBAKEL, Johannes M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
ENCODING A RIPENING FORM OF A POLYPEPTIDE HAVING
TITLE OF INVENTION: RHANNOGALACTURONASE ACTIVITY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABRY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.

CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/536,150
FILING DATE: 29-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/061,062
FILING DATE: 14 MAY 1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 202390/R 7262 (V)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-536-150-6

Query Match 10.1%; Score 241; DB 3; Length 440;
Best Local Similarity 24.0%; Pred. No. 2e-13;
Matches 88; Conservative 61; Mismatches 166; Indels 32; Gaps 11;

QY 70 KNGIKVINIVSFGAKGDKTYDNIAPFOAMNEACSSRTPOVFVPPKKNKNTLLKQITSGP 129
DB 33 KGATKTCNIIISYGAVANSTDVGPALISAW-AAKSGGLV--YIPSGNYALNTWVTLTG 89
QY 130 CRSSISVK--IFGSLASSKISDYKDRRLMTAFPSVONLVVGGGNTNGNQVWMPSSCK 187
DB 90 SATAIQDGLIYRTGTASGNMIAVLTDTTDFELFSTSK-----GAVOGFGYVYHAEQTY 143
QY 188 INKSLPRDAPALTFNCKLKNVNLKSKNAQOIHKFESCTNVVASNIMINSAKSPN 247
DB 144 GARILRLTDV-----THFSVHDVILVDAPAFHFTMTDCSDGEVYNNALIRGNBG-G 193
QY 248 TDGVHVSNTQYIOISDTTIGTDDCISIVSGSQNVQATNITCGPHGISIGSLGSGNSEA 307
DB 194 LDGIDVWGSN-IWHDVEVTNKDECVTKSPANNILVESIYCNMSSGCGAMGSLG---ADT 249
QY 308 YVSNVTNNEAKITGAENGVRKITWQSGSQASNIKEFLNEMQDVKYPIIIDONYCDRVER 367
DB 250 DVTDIVRNVYTWSSNQMYIKS-NGSGTVSNVLTLENFIGHGNAVSLDIDIGWSSMTAV 308
QY 368 CIOQFSAVOVKNVYENIKGT---SATKVAIKPDCSTNFPCEGIMENINLVGSGKPS 424
DB 309 A---GDGVQLNNTIVKWKGTGANGATRPPIRVVCSDTAPCTDLTLDIAIWTSGSSEL 365
QY 425 ATCKNVH 431
DB 366 YLCRSAY 372

RESULT 15
US-08-536-150-8
Sequence 8, Application US/08536150
Patent No. 6013489
GENERAL INFORMATION:
APPLICANT: MUSTERS, WOUTER
APPLICANT: STAM, HEIN
APPLICANT: SUYKERBOYK, MARIA E.
APPLICANT: VISSER, JACOB

APPLICANT: VERBAKEL, Johannes M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
TITLE OF INVENTION: ENCODING A RIPENING FORM OF A POLYPEPTIDE HAVING
TITLE OF INVENTION: RHANNOGALACTURONASE ACTIVITY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARRY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/536,150
FILING DATE: 29-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/061,062
FILING DATE: 14 MAY 1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 202390/R 7262 (V)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-536-150-8

Query Match 10.1%; Score 241; DB 3; Length 440;
Best Local Similarity 24.0%; Pred. No. 2e-13;
Matches 88; Conservative 61; Mismatches 166; Indels 32; Gaps 11;

QY 70 KNGIKVINIVSFGAKGDKTYDNIAPFOAMNEACSSRTPOVFVPPKKNKNTLLKQITSGP 129
DB 33 KGATKTCNIIISYGAVANSTDVGPALISAW-AAKSGGLV--YIPSGNYALNTWVTLTG 89
QY 130 CRSSISVK--IFGSLASSKISDYKDRRLMTAFPSVONLVVGGGNTNGNQVWMPSSCK 187
DB 90 SATAIQDGLIYRTGTASGNMIAVLTDTTDFELFSTSK-----GAVOGFGYVYHAEQTY 143
QY 188 INKSLPRDAPALTFNCKLKNVNLKSKNAQOIHKFESCTNVVASNIMINSAKSPN 247
DB 144 GARILRLTDV-----THFSVHDVILVDAPAFHFTMTDCSDGEVYNNALIRGNBG-G 193
QY 248 TDGVHVSNTQYIOISDTTIGTDDCISIVSGSQNVQATNITCGPHGISIGSLGSGNSEA 307
DB 194 LDGIDVWGSN-IWHDVEVTNKDECVTKSPANNILVESIYCNMSSGCGAMGSLG---ADT 249
QY 308 YVSNVTNNEAKITGAENGVRKITWQSGSQASNIKEFLNEMQDVKYPIIIDONYCDRVER 367
DB 250 DVTDIVRNVYTWSSNQMYIKS-NGSGTVSNVLTLENFIGHGNAVSLDIDIGWSSMTAV 308
QY 368 CIOQFSAVOVKNVYENIKGT---SATKVAIKPDCSTNFPCEGIMENINLVGSGKPS 424
DB 309 A---GDGVQLNNTIVKWKGTGANGATRPPIRVVCSDTAPCTDLTLDIAIWTSGSSEL 365
QY 425 ATCKNVH 431
DB 366 YLCRSAY 372

Search completed: July 21, 2004, 17:09:27

✓ Tue Jul 27 08:55:41 2004

us-10-018-604-2.ra1

Page 9

Job time : 29.379 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:08:27 ; Search time 79.2655 Seconds
(without alignments)
1802.048 Million cell updates/sec

Title: US-10-018-604-2

Perfect score: 2390
Sequence: 1 MWIQNSILLIILIFASSIS.....VTPHCTSLSEIDELALNY 457Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1285356 seqs, 312560742 residues

Total number of hits satisfying chosen parameters: 1285356

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Published_Applications_AA.*
2: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1113	46.6	433	13	US-10-151-668-2
2	1011	42.3	393	16	US-10-437-963-127090
3	1010	42.3	452	16	US-10-437-963-127091
4	988.5	41.4	444	16	US-10-425-963-111044
5	870	36.4	319	12	US-10-425-114-69361
6	808	33.8	508	16	US-10-437-963-154485
7	787.5	32.9	514	10	US-09-847-208-69
8	785	32.8	534	16	US-10-437-963-134257
9	739	30.7	485	16	US-10-437-963-175508
10	734	30.7	503	16	US-10-437-963-172689
11	728	30.5	443	12	US-10-424-599-235443
12	728	30.5	456	12	US-10-425-114-44707
13	724.5	30.3	398	12	US-10-362-091-4
14	723.5	30.3	492	12	US-10-424-599-284649
15	701.5	29.4	400	16	US-10-437-963-155040

16	700	29.3	573	12	US-10-425-114-43413	Sequence 43413, A
17	699	29.2	496	12	US-10-425-114-66151	Sequence 66151, A
18	682	28.5	571	16	US-10-437-963-159100	Sequence 159100, A
19	681	28.5	737	16	US-10-437-963-162350	Sequence 162350, A
20	678	28.4	457	16	US-10-437-963-146166	Sequence 146166, A
21	666	27.9	393	16	US-10-437-963-116114	Sequence 116114, A
22	664.5	27.8	415	12	US-10-424-599-238512	Sequence 238512, A
23	660	27.6	422	12	US-10-424-599-176330	Sequence 176330, A
24	653.5	27.3	385	12	US-10-425-114-41984	Sequence 41984, A
25	653.5	27.3	411	12	US-10-425-114-38472	Sequence 38472, A
26	653.5	27.3	445	12	US-10-425-114-51531	Sequence 51531, A
27	649.5	27.2	449	16	US-10-437-963-110174	Sequence 110174, A
28	643.5	26.9	602	16	US-10-437-963-127089	Sequence 127089, A
29	642	26.9	320	12	US-10-424-599-204806	Sequence 204806, A
30	641.5	26.8	544	16	US-10-427-963-110126	Sequence 110126, A
31	635.5	26.6	539	12	US-10-424-599-176466	Sequence 176466, A
32	632	26.4	386	12	US-10-425-114-50991	Sequence 50991, A
33	630	26.4	410	12	US-10-425-114-38475	Sequence 38475, A
34	630	26.4	414	12	US-10-425-114-62958	Sequence 62958, A
35	630	26.4	414	12	US-10-425-114-64531	Sequence 64531, A
36	627	26.2	407	12	US-10-425-114-65565	Sequence 65565, A
37	627	26.2	410	12	US-10-425-114-46486	Sequence 46486, A
38	627	26.2	412	12	US-10-425-114-46801	Sequence 46801, A
39	627	26.2	413	12	US-10-425-114-52982	Sequence 52982, A
40	627	26.2	413	12	US-10-425-114-56708	Sequence 56708, A
41	627	26.2	415	12	US-10-425-114-46800	Sequence 46800, A
42	627	26.2	415	12	US-10-425-114-54673	Sequence 54673, A
43	627	26.2	415	12	US-10-425-114-64510	Sequence 64510, A
44	627	26.2	416	12	US-10-425-114-65564	Sequence 65564, A
45	627	26.2	416	12	US-10-425-114-72785	Sequence 72785, A

ALIGNMENTS

RESULT 1
US-10-151-668-2
; Sequence 2, Application US/10151668
; Publication No. US20020184660A1
; GENERAL INFORMATION:
; APPLICANT: UMYSKOV, Peter
; APPLICANT: CHILD, Robin
; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els
; APPLICANT: BORKHARDT, Bernard
; APPLICANT: SANDER, Lilli
; APPLICANT: PETERSEN, Morten
; APPLICANT: BONDGARD, Poulsen, Gert
; APPLICANT: BOTTERMAN, Johan
; TITLE OF INVENTION: Seed Shattering
; FILE REFERENCE: 2121-0138P
; CURRENT APPLICATION NUMBER: US/10/151,668
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US/09/051,239
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: PCT/EP96/04313
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: EP 95 402241.4
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: EP 95 203328.0
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Strain cv. Topaz.
US-10-151-668-2
Query Match 46.6%; Score 1113; DB 13; Length 433;
Best Local Similarity 49.9%; Pred. No. 1.8e-97;

[illegible]

```

QY      310 SNVNVENKIIIGAENGVAIKITMOGSGGASNIKFLENVMODKVPIIIDONYCDRVEPCI 369
          |::: :: |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      181 SGIITDSQLHGHTNGARIKRTYQGSSGYAKDITTFOMVMVDVKNPFIIDONYCDRAKPCG 240
          |::: ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      370 QQSFAVVQKNVVENINIKGTSAITKYAIKFDCTNPFCEGIIMENINLVGSCK-PSEATCK 428
          |::: ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      241 EGSAAVGVGVFKNIRGSTSTKDALKMNCSENVPQGITLIDILKMDDKGNTIRSTCQ 300
          |::: ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      429 NVHNNNAEHVTPE-CTSLE 446
          |::: ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      301 NAKWTEFGTVRPQPCTAIRK 319
          |::: ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 6
US-10-437-963-154485
; Sequence 134485, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OR INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154485
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54340C.1.pep
US-10-437-963-154485

Query Match      33.8%   Score 808; DB 16; Length 508;
Best Local Similarity 36.4%; Pred No.3,6e-68;
Matches 173; Conservative 91; Mismatches 167; Indels 44; Gaps 7;
```

```

QY      8 ILLIIIFASSISTCRSNVIDNLFKQVYVNLBEOFAHDFQAYL-----SY 54
D      18 VLVALMTYIMAAKTASGGGGDGGAGGHD-----HDHFQPLKIMNGGADAKEDY 69
QY      55 LSKNIESNNIDKVDKN-----GIKVINVLSPGAKDGTYNIAEQ 97
D      70 LNMDDDDDDDEDEEEEAQVMAAKCRPPAGRVNVNDVSFGAAGDGCSDTEATLN 129
QY      98 AMNEACSRTPVQVVPKKNKYLLKQITFSGPCRSSISVKIFGSLSEASKISDY--KRR 155
D      130 AMKRACSLNNAV-FLVPGRRYKVGAAAFISPCKNRMITIQGTTVADESEMDPASPR 188
QY      156 LMIAFDSYQNLVVGGGGTINGNGQVWMPSSCKINKSLPCRPDAPALTFFMCKNLKVNLT 215
D      189 LMLFSLGAGARIQGGGLIDGSGSKMANSCIKIDRSKCKAPALITIDSCRGVSRLR 248
QY      216 SKNAQOIHKEESCCTNVVSNLMNINAKSPNTDGVHVSNTQYIQISDTIIGTDDCISI 275
D      249 LQNAQOMHLLVRSRSDVRLASVRVDSPEDSPTDGIHVADSTAVTIQSCRITATGDDCISI 308
QY      276 VSGSQNVQATNITGPGHGISIGSGNSEAVYSNVYNNAKIIIGENGVRITMOGGS 335
D      309 SNGSFAPVRMRIDCGPFGHISIGSGGGAFAVDGSLDGAFAVAQNGVRITMOGGA 368
QY      336 GQASNIKFLNVEMQDVKKYPIIIDQNYCDRVEPCIQSFSAVQKVVYENIKGTSATKVAI 395
D      369 GYVNRVAFAGRVGVGDHPIVIDQFYCDATPCRNRTSNVAVSGVFRNITGTAARAIAI 428
QY      396 KFDGCTNPPCEGIIEMENINLVGSGKSEATCKN--VHFNNAEHVTHPCHSLEIS 448
D      429 RLACSDAVPCVGIYLSIDILREDDGGGVQVTCNCAMGFDDG-RVSPADCLRTS 482

```

RESULT 7

```

US-09-847-208-69
; Sequence 69, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: UC67.002A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: Fasteq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Cryptomeria japonica (Japanese cedar)
US-09-847-208-69

```

```

Query Match      32.9%; Score 787.5; DB 10; Length 514;
Best Local Similarity 38.5%; Pred. No. 3.3e-66;
Matches 173; Conservative 78; Mismatches 163; Indels 35; Gaps 9;

QY      8 ILLIIIFASSISTCRSNVIDNLFKQVYVNLBEOFAHDFQAYLSTYSKNTESNNNDK 67
D      12 VAMOLITMAA-----EDQSAQIMLSDIEQ-----YL-----RSNRSIRK 47
QY      68 VD---KNGIKYINVLFGAKGDKGTYNIAFEQAMNEACSRTPVQVFPKKNKYLLKQI 124
D      48 VHSRRHAINIFNEVEKIGAVDEGHDCTEAFSTAMQAC-KKPSAMLLVPSNKKFVVNLT 106
QY      125 TFGSPCRSSISVKIFGSLSEASKISDYKDRRLMTAFDSVQNLVVGGGGTINGNGQVWMP 184
D      107 FPNGPQPHFTFKYDGIILAAQVNPASMKNNRIWLFQAKLTGFTLMGKGVIDQGRKQWAG 166
QY      185 SCR-INKSLPC--RDATALTFFMCKNLKVNLSKNAQOIHKEESCCTNVVSNLMNINA 241
D      167 QCKWVNGREICNDRPETAIKFDFSTGLIIGLKLXNSPEPHLVFGNCEGVKIIIGISITA 226

```

```

QY      242 SAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITGPGHGISIGSLG 301
D      227 PRDSPNTDIDIPASKKNHLPKTLTGTDCAVAGTSSNIVIELIGPFGHGISIGSLG 286
QY      302 SGNSEAVYSNVTVNEAKIIGAENGVRITMOGSGQASNIKFLNVEMQDVKKYPIIIDQNY 361
D      287 RENSRAEVSYVHVAGAKRIDQNGRLIKTWOGSGMAHIIYENEMINSENPILINQFY 346
QY      362 CDRVEPCIQFSANVQKVVYENIKGTSATKVAIKFDGCTNPPCEGIIEMENINLVGSGK 421
D      347 CTSASACQMSRAVQIIDVYTKNIRGTSATMAAIQLKCSDEMPCKDILSLKLTSGK 406
QY      422 PSBATCKNVHFN--AEHVTCHSLEIS 448
D      407 --TASCLNDNANGYFSGHVTPACKNLSFS 433

```

RESULT 8

```

US-10-437-963-134257
; Sequence 134257, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bonkharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 134257
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_36049C.1.pep
US-10-437-963-134257

```

```

Query Match      32.8%; Score 785; DB 16; Length 534;
Best Local Similarity 41.1%; Pred. No. 6.1e-66;
Matches 157; Conservative 57; Mismatches 146; Indels 22; Gaps 4;

QY      81 FGAKGDKGTYNIAFEQAMNEACSRTPVQVFPKKNKYLLKQITFSGPCRSSISVKIFG 140
D      130 FGAVDGTINDYQFLDMNAACASTEPAYLAIVAGTGYQWPRALAPCKKTLKMTLSG 189
QY      141 SLEASKSISDY--KDRRLMIAFDSVQNLVVGGGGTINGNGQVWMPSSCKINKSLPCR--D 196
D      190 TIAPASPDEWAGBDPKMLIVFRVDDLVSGGGTIDMGAEWMAWSCCKRKYPCSTVS 249
QY      197 APTALTFFMCKNLKVNLSKNAQOIHKEESCCTNVVSNLMNINAKSPNTDGVHVSNT 256
D      250 APKALOEEBCRRVAVQITWONGPQFLMFRCDVNASFLRVVAPESSPTDGIHLNLT 309
QY      257 QYIQISDTIIGTDDCISIVSGSQNVQATNITGPGHGISIGSGNSEAVYSNVYNNAKIIIG 316
D      310 THAQIMNLSTGDDCYSWNGNCSDVAYKDISCPGHSISIGSGKRTTIRIENAVDT 369
QY      317 AKIIGAENGVRITMOGSGQASNIKFLNVEMQDVKKYPIIIDQNYCDRVEPCIQO----- 371
D      370 CLTNTNTNGVRIRKQMGGMGYAHLRFEGIVMKVNSPIIIDYXCDQPTPCANOSTSS 429
QY      372 -----FS-----AVQKVVYENIKGTSATKVAIKFDGCTNPPCEGIIEMENINLVGE 418
D      430 SLKLGNSSLIMQOAVAEVRKIEFAGIRGTSATQAIKLACSDAVPCPDELRLRVNLTMV 489

```

QY	419	SGKPS	SEATCK	VHNN	AHE	VTP	440
			:				
Db	490	GGASA	FCHRA	SGKA	GA	VVP	511

RESULT 9

```

US-10-437-963-175508
; Sequence 175508, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 175508
; LENGTH: 485
; TYPE: PRF
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73346C.1.pep
US-10-437-963-175508

```

Query Match	30.9%	Score 739;	DB 16;	Length 485;
Best Local Similarity	42.6%	Pred. No. 1.3e-61;		
Matches 153;	Conservative 62;	Mismatches 128;	Indels 16;	Gaps 7;

```

OY      75  VJNVJSPGAKGKGYKXNDIAFOQANNEGCSSTPQVFWKKXNVLKQIFSGGCRSI 134
Db      89  VEDPARFAVAGDGTDDTEAPRAAMRACAVESAV-IVSPSGTFTITTTTFTGCKRGL 147
OY      135  SVKJFGSL--EASKISDYKDRIMIAFDVSQNLVVGGGGTINGGCVWMPDSCKINKS 191
Db      148  VFQYDGVGLMPPGDPCWPPSDNRROMLTFVSNLDGLTLRGAGTIBNGGGMWMLPKCKRHG 207
OY      192  L-----PCRDAPALFTFMNCKNKXNNLKSKNAQOIHIXEESCCTNVASLMTMASAK 244
Db      208  ENGSTLRRGC-DSPLLVRFEMSRNLVBEGLRKNENSPERHFRFDGGSYARVDELSTSPAN 266
OY      245  SPNTDGVHVSNTQYIQTSDTIIGTDDDCISIVSGSQNVQANITGPGHGISISGLSGN 304
Db      267  SPNTDGIHVENTQRAIVNYSMTSNCDDCISICTGSDYDIDQNVSGPBGHGISISGLGYHN 326
OY      305  SEAYVSNVTYVNAKIIIGANGVRITKMOGSGQANIKELNEMQDVKYPIIIDQNYC-D 363
Db      327  SQACVANTYVRAVYLRNSDNGRLRTKMOGMSVSGINFDIYMSVENVRCIIITDYCYLD 386
OY      364  RVEPCIQGFSAVQVKNVYENIKGTSATKVA-1KEDCSINPCEGIINENINLVGSESG 421
Db      387  K--RCMNDSTAVHVDVSYANVRGGDYVAAIIFHACGSPVCTNITMSEVILLFFSGG 443

```

RESULT 10
ITE=10-427

US-10-447-963-172689
 ; Sequence 172689, Application US/104379633
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Bird

```

: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53221)B
: CURRENT APPLICATION NUMBER: US/10/437,963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 172689
: LENGTH: 503
: TYPE: PRT
: ORGANISM: Oryza sativa
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT4530_70800C.1 pep
: US-10-437-963-172689

```

Query Match	Score 734;	DB 16;	Length 503;
Best Local Similarity	40.7%;	Pred. No. 4.2e-61;	

QY	78	LLSTGAGDGGTYNIMFEQMANMGASRTFVQVJVPVKNKNXYLLKQITFSGPCSSISVX	137
Db	114	VKDFGAGDGVTDITDMLKTRMTDRCACDDGAGVVLAAAGRSFLIHITTVTFSPCGSVTLQ	173
QY	138	IFGSLIEMSSKSIDY--KDRRLMIAFDVQNLVVGGGGTINGNGVWVWSSCKINKS----	191
Db	174	VDGTTIVAPSEPALTMPANKRRWMLVFYRADGVLGAGLIDQGGKXWMDLPCKPHKGNTH	233
QY	192	LPCCDAPLALTFWNCKNLKVNNLKSKAQOIHIKEFESTVNVASNLIMNASAKSPNTDGV	251
Db	234	GPC-DSPAFAMFAISNNVTVFGLKVQMSPEPFHFEFNDNCNGRVLDGLSISSPALSPNTDGI	292
QY	252	HVSNMTOYIQISDITLIGTGDCCISTVSGSQNQANITCGPHGHSISLSGSGNSEAYVSN	311
Db	293	HVENTSDVLTNTVASSGGDCVSIAGAGLNVHIENTVCGPHGHSISLSGAGATKACVYN	352
QY	312	VTVNEAKIGAEENVRIKTKWQSGSGSOASNIKFLNEVQCVKXPPIIIDQNYCDRAVEPCIQ	371
Db	353	VTVNAVAIRHSDNGVRIKTKWGGSGGSVAFAFENVRRDAVARNPPIIIDQYVC-LKSCENE	411
QY	372	FSAVQVKNVYENIKGTSATK-VAIKEDCSTNFPCEGIIMENINLVGBSGKP-SEATCKN	429
Db	412	TTAFVANGVSSGIRGTVDVGRPIIHFGCSAPVAPCTNITLSDVELLPASGDTVDPEFCWM	471
QY	430	VHFNNAEHVTPHCISL 445	
Db	472	VYGNAAFPVPPVSC L 487	

RESULT 11

```

US-10-424-599-235443
; Sequence 235443, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21 (53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 235443
; LENGTH: 443
; TYPER: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_54632C.1.pep
US-10-424-599-235443

```

Query Match	30.5%;	Score 728;	DB 12;	Length 443
Best Local Similarity	42.9%;	Pred. No. 1.3e-60;		


```
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 284649
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_99063C.1.pep
US-10-424-599-284649
```

Query Match 30.3%; Score 723.5; DB 12; Length 492;

Best Local Similarity 39.6%; Pred. No. 4.1e-60; Mismatches 157; Indels 15; Gaps 7;

```
Matches 155; Conservative 64; Mismatches 157; Indels 15; Gaps 7;
QY 75 VINVLSEFGAKGDKTYDNTAFEOAMNEACSSRTPOFVVPKKNYLLKQITPSGCRSS1 134
DB 96 VEDVSSFGAVGDCADDTFRAPRAAMRAACAVDSGI-VLAPENVYFKITSTIFSGPCKPGL 154
QY 135 SVKIRGSEASSKISIDYKD--RLMLIADSVQNLVVGGGGTNGQVWMPSSCKLNK- 190
DB 155 VFOVDLTLVAPDGPNSWPEADSRNQWLVEYRLDQMLNGLTLEGSGDMWDLPCKEPHNG 214
QY 191 -----SLPCRDAPLTALTFWNCCKNLKYNLKSNAQOIHIFESCTNVVASNLMINASAK 244
DB 215 PNGKITLSGFC-GSPAMIRRFMSNLVKYKGLKTONSGQFMIFNGCGVLIDKLISSEPKL 273
QY 245 SPNTDGVHVSNTQYIQISDTITGTGDDCISIVSGSONVQATNTTCGPHGISIGSLGSGN 304
DB 274 SPNTDGIHYENSGLRKYVGIYNSMISNGDDCISIGPSSNVDIAGLTGPHSGISIGSLGVAN 333
QY 305 SEAYVSNVTVNNAKTIIGAENGVRITKTWOGSGGASIKFLANVMQVYKPIIIDQNYCR 364
DB 334 SQACVSNLTVRDSIIRRESNGRLIKITWOGMGSSVSLIRENIOENEVNGCIITIDQYCCIS 393
QY 365 VEPICIQFSAVOVKNVVYENIKGTSATKYA-IFDCSTNPPCEGIIEMENINLVG-ESGKP 422
DB 394 KE-CLNQTSAVHVNDVSYNIKGTVDYRTAPIHFACSDTVACTNTITLSEVELPFEGALL 452
QY 423 SEATCKNVHFNNAEHVTPHCTSLSESEAL 453
DB 453 DDPFCWNAVGTQETLTIPINCLREGDPETV 483
```

RESULT 15

US-10-437-963-155040

; Sequence 155040, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Bouharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 155040

```
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54842C.1.pep
US-10-437-963-155040
```

Query Match 29.4%; Score 701.5; DB 16; Length 400;

Best Local Similarity 40.1%; Pred. No. 3.9e-58; Mismatches 148; Conservative 69; Mismatches 133; Indels 19; Gaps 8;

```
QY 72 GIKVINVLSEFGAKGDKTYDNTAFEOAMNEACSSRTPOFVVPKKNYLLKQITPSGCR 131
DB 25 GAAYVNESYDARDPGRITDASRALASAWGAACRSEPPATVYVDDGE-FVSHSAFAGPCS 83
QY 132 SS-ISVKIRGSLDEA--SSKISDYKDRRLMLIADSVQNLVVGGGGTNGQVWMPSSCKIN 169
DB 84 GGRMTVQIDGTLVAPDGPNSWPEADSRNQWLVEYRLDQMLNGLTLEGSGDMWDLPCKEPH 140
QY 190 KSLPCRDAPLTALTFWNCCKNLKYNLKSNAQOIHIFESCTNVVASNLMINASAKSPNTD 249
DB 141 GHGCGCPDGAITSKKNVNSRDVVISGVKSANSELVHVIDCGEYAVQDARIVAPGSSPNTD 200
QY 250 GVHVSNTQYIQISDTITGTGDDCISIVSGSONVQATNTTCGPHGISIGSLGSGN 309
DB 201 GIHVOSSSAVTITGASIGTGDCCISIVGPSTNMLRVHVS CGPHGISIGSLGSEEGGV 260
QY 310 SVTVNNAKTIIGAENGVRITKTWOGSGGAS-----NIKFLANVMQVYKPIIIDQNYC 362
DB 261 ENVTVSGAATVGTENGLRIKTN---GRARSGAYRVGVFEBALMRDVSNPILIIDQSYC 316
QY 363 --DVEPCIQFSAVOVKNVVYENIKGTSATKYAIFDCSTNPPCEGIIEMENINLVGSG 420
DB 317 PNDGGGCGCHQSSDVQISGVTYTDIIGSSASQVAVFDCSASRPGCGGLGIQDILKLPDGG 376
QY 421 KPSEATCKN 429
DB 377 KPSEATCKN 385
```

Search completed: July 21, 2004, 17:17:48

Job time : 82.2655 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:04:03 ; Search time 29.3576 Seconds
(without alignments)
1497.381 Million cell updates/sec

Title: US-10-018-604-2

Perfect score: 2390
Sequence: 1 MWIQNSILLIIRFASIS.....VTPHCTSLRISDEALVNY 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2390	100.0	457	2 A25534	polygalacturonase
2	1142	47.8	462	1 S31195	polygalacturonase
3	1128	47.2	461	2 T08215	polygalacturonase
4	1127	47.2	431	2 T46187	polygalacturonase
5	1115.5	46.7	426	2 H84846	probable polygalac
6	1108	46.4	458	2 S71523	polygalacturonase
7	1091	45.6	460	2 T17011	polygalacturonase
8	898.5	37.6	459	2 D96833	hypothetical prote
9	840.5	35.2	468	2 H96728	probable polygalac
10	796.5	33.3	514	2 UC7100	polygalacturonase
11	787.5	32.9	434	2 A96609	probable polygalac
12	787.5	32.9	514	2 S48730	Cry j II protein -
13	764	32.9	514	2 JC2498	second major aller
14	764	32.0	507	2 JC7366	Jun a 2 protein -
15	749.5	31.4	1161	2 B86368	protein F28C11.9 l
16	741.5	31.0	491	2 B86155	probable polygalac
17	736	30.8	423	2 T08213	polygalacturonase
18	728	30.5	405	2 D84871	probable polygalac
19	727	30.4	392	2 S57806	polygalacturonase
20	718.5	30.1	395	2 T05906	probable polygalac
21	718	30.0	435	2 T48638	polygalacturonase
22	715	29.9	392	2 T00669	probable polygalac
23	711.5	29.8	387	2 T04320	polygalacturonase
24	710.5	29.7	365	2 T04319	polygalacturonase
25	708.5	29.6	422	2 S34199	polygalacturonase
26	704.5	29.5	387	2 T04322	polygalacturonase
27	704	29.5	374	2 B85421	probable polygalac
28	704	29.5	374	2 T04672	polygalacturonase
29	702.5	29.4	394	2 G86190	hypothetical prote

30	701.5	29.4	542	2 H86239	protein F20B24.8 l
31	698.5	29.2	452	2 C85024	probable polygalac
32	696	29.1	384	2 E84871	probable polygalac
33	690	28.9	540	2 B96631	probable polygalac
34	688.5	28.8	492	2 C96521	protein F21D18.18
35	687.5	28.8	393	2 S40123	polygalacturonase
36	687	28.7	383	2 T47809	polygalacturonase
37	684	28.6	397	2 F96680	F5114.10 [Imported
38	683	28.6	394	2 F86190	hypothetical prote
39	680.5	28.5	394	2 T00668	probable polygalac
40	678.5	28.4	445	2 S34266	polygalacturonase
41	673.5	28.2	387	2 T07591	polygalacturonase
42	673	28.2	444	2 S34200	polygalacturonase
43	667	27.9	407	2 S52006	polygalacturonase
44	662.5	27.7	664	2 A84742	probable polygalac
45	659	27.6	414	2 E85204	polygalacturonase-

ALIGNMENTS

RESULT 1

A25534 polygalacturonase (EC 3.2.1.15) precursor - tomato

N:Alternate names: pectinase; poly[1,4-alpha-D-galacturonide]glucanohydrolase
C:Species: Lycopersicon esculentum (tomato)

C>Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #ext_change 22-Jun-1999
C:Accession: A25534; J0156; S06340; S02101; S08504

R:Grierson, D.; Tucker, G.A.; Keen, J.; Ray, J.; Bird, C.R.; Schuch, W.
Nucleic Acids Res. 14, 8595-8603, 1986

A>Title: Sequencing and identification of a cDNA clone for tomato polygalacturonase.
A:Reference number: A25534; MUID:87066731; PMID:3786135

A:Accession: A25534
A:Molecule type: mRNA

A:Residues: 1-457 <GRT>

A:Cross-references: GB:X04583; NID:g19291; PIDN:CAA28254.1; PID:g19292
R:Bellappa, D.; Bennett, A.B.

Plant Physiol. 86, 1057-1063, 1988
A>Title: In vitro synthesis and processing of tomato fruit polygalacturonase.

A:Reference number: J0156
A:Accession: J0156

A:Molecule type: mRNA

A:Residues: 1-115

A:Cross-references: GB:M20269; NID:g170470; PIDN:AAA34177.1; PID:g170471
A:Experimental source: fruit

R:Sheehy, R.E.; Pearson, T.; Brady, C.J.; Hiatt, W.R.
Mol. Gen. Genet. 208, 30-36, 1987

A>Title: Molecular characterization of tomato fruit polygalacturonase.
A:Reference number: S06340
A:Accession: S06340

A:Molecule type: mRNA

A:Residues: 1-457 <SHE>

A:Cross-references: EMBL:X05556; NID:g19297; PIDN:CAA29148.1; PID:g19298
A>Note: part of this sequence, including the amino end of the mature protein, was confir

R:Rose, R.E.; Houck, C.M.; Monson, E.K.; Dejeus, C.E.; Sheehy, R.E.; Hiatt, W.R.
Nucleic Acids Res. 16, 7191, 1988

A>Title: The nucleotide sequence of the 5' flanking region of a tomato polygalacturonase
A:Reference number: S02101; MUID:88303350; PMID:3405769

A:Accession: S02101
A:Molecule type: DNA

A:Residues: 1-93 <ROS>

A:Cross-references: EMBL:X07410; NID:g19295; PIDN:CAA30308.1; PID:g19296
R:Bird, C.R.; Smith, C.J.S.; Ray, J.A.; Mourreau, P.; Bevan, M.W.; Bird, A.S.; Hughes, S

Plant Mol. Biol. 11, 651-662, 1988
A>Title: The tomato polygalacturonase gene and ripening-specific expression in transgen

A:Reference number: S08504
A:Accession: S08504

A:Molecule type: DNA

A:Residues: 1-457

A:Cross-references: EMBL:X14074; NID:g19305; PIDN:CAA32235.1; PID:g295813
C:Comment: The polygalacturonase, a single cell wall enzyme, is the major enzyme mediat

C:Superfamily: polygalacturonase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-71/Domain: propeptide #status predicted <PRO>
F:72-444/Product: polygalacturonase 2A #status experimental <MAT>

Query Match 100.0%; Score 2390; DB 2; Length 457;

Best Local Similarity 100.0%; Pred. No. 9,6e-162;

Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MVIQRNSILLIIIFASISITCRSNVIDNLFKQVYNNILIEQFAHDFQAYLSTLKNIE 60
DB 1 MVIQRNSILLIIIFASISITCRSNVIDNLFKQVYNNILIEQFAHDFQAYLSTLKNIE 60
QY 61 SNNNIDKVDKNGKIVINVLSEFGAKGDKGTNDNIAFEQANNEACSSRPVQFVVKNNY 120
DB 61 SNNNIDKVDKNGKIVINVLSEFGAKGDKGTNDNIAFEQANNEACSSRPVQFVVKNNY 120
QY 121 LKQITFSGPCRSSISVYKIFGSLSEASSKISDYKORRLMIAFDSVQNLVVGSGGTINGQV 180
DB 121 LKQITFSGPCRSSISVYKIFGSLSEASSKISDYKORRLMIAFDSVQNLVVGSGGTINGQV 180
QY 181 WWPSSCKINSLPCRDAPLTATFWNCKNLKVNNLKSNAAQOIHFKESCTNVVASNLMIN 240
DB 181 WWPSSCKINSLPCRDAPLTATFWNCKNLKVNNLKSNAAQOIHFKESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHSNTQYIQTSDTIIGTGDDCISIVSGSQNVQATNITCGPHGISIGSL 300
DB 241 ASAKSPNTDGVHSNTQYIQTSDTIIGTGDDCISIVSGSQNVQATNITCGPHGISIGSL 300
QY 301 GSGNSEAYVSNVTYNEAKIIGAENGVRKKTWQSGGQASNIKFLINVMQVYPIIIDQN 360
DB 301 GSGNSEAYVSNVTYNEAKIIGAENGVRKKTWQSGGQASNIKFLINVMQVYPIIIDQN 360
QY 361 YCQRVEPCIOQFQAVQKVNVTENIKGTSATKYAIKRDCSTNPFCEGIIEMENINLVGESG 420
DB 361 YCQRVEPCIOQFQAVQKVNVTENIKGTSATKYAIKRDCSTNPFCEGIIEMENINLVGESG 420
QY 421 KPSEATCKNVHPNNAEHVTPHCTSLSEISDEBALLYNY 457
DB 421 KPSEATCKNVHPNNAEHVTPHCTSLSEISDEBALLYNY 457
```

RESULT 2

S31195
polygalacturonase (EC 3.2.1.15) - avocado
C/Species: Persea americana (avocado)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: S31195; S28072
R/Dopico, B.; Lowe, A.L.; Wilson, I.D.; Merodio, C.; Grierson, D.
Plant Mol. Biol. 21, 437-449, 1993
A>Title: Cloning and characterization of avocado fruit mRNAs and their expression during
A/Reference number: S31195; MUID:93184201; PMID:8095163
A/Accession: S31195
A:Molecule type: mRNA
A:Residues: 1-462 <DOP>
A/Cross-references: EMBL:X66426
R/Dopico, B.
submitted to the EMBL Data Library, May 1992
A/Reference number: S28072
A/Accession: S28072
A:Molecule type: mRNA
A:Residues: 1-181, 'S', 183-462 <DO2>
A/Cross-references: EMBL:X66426; NID:g22630; PIDN:CAA47055.1; PID:g22631
C/Superfamily: polygalacturonase
C/Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 47.8%; Score 1142; DB 1; Length 462;

Best Local Similarity 55.3%; Pred. No. 2,6e-73;

Matches 224; Conservative 67; Mismatches 106; Indels 8; Gaps 5;

```
QY 28 DNNLFKQVYNNILIEQFAHDFQAYLSTLKNIESNNNIDKVDKNGKIVINVLSEFGAKG 87
DB 40 DQRAYPTVFGPIID-EPSSIMGFEPSITLS--LFRFNVGGEPSPTDITSDVDFGARGDG 96
```

QY 88 KTYDNIAFEQANNEACSSRPVQFVVKNNYLLKQITFSGPCRSSISVYKIFGSLSEASSK 147
DB 97 -TDTYAFEFKAMKDACSSGVL--IYENKYNLLKQITFSGPCSSDKRLKRTTEASSD 153

QY 148 ISDY--KDRRLMIAFDSVQNLVVGSGGTINGQVWPPSSCKINKSLPCRDAPLTATFWN 205
DB 154 QSDWVGHNRRKRTIEFEDISMLTEGGGTINGNGETWDDSCSKRKSIPCKSAPALTFRS 213

QY 206 CKNLKVNNLKSNAAQOIHFKESCTNVVASNLMINAKSPNTDGVHSNTQYIQTSDTI 265
DB 214 CKNLIVSDLSIKOSQKMHLSFDCKQDYASNLMVLAPEHSPNTDGHITGTQIHWMSV 273

QY 266 IGTGDDCISIVSGSQNVQATNITCGPHGISIGSLGSGNSEAYVSNVTYNEAKIIGAENG 325
DB 274 IGTGDDCISIESGSKVYLAINTITCGPHGISIGSLGGRNEAHVSGVLVDGKNLFTTNG 333

QY 326 VRIKTWQSGGQASNIKFLINVMQVYPIIIDQNTCDRAVEPCIQPSANQKVNVTENI 385
DB 334 IRIKTWQSGGSAKNIFQVNIWENVTNPIIIDQYCDSDKDPCEQESAVVAVSNVANI 393

QY 386 KGTSAATKVAIKPDCSTNPFCEGIIEMENINLVGSGKPSSEATCKVY 430
DB 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

QY 394 RGTSASEVAVKFDCKSSPCQGIYVGNINLVGNGKRTTWSCSNI 438

```
Db 313 AGKSEAENVNVVDPAKESGTSNGVRIKTMQGGKGAQNIIFQNIWMNVNPIIINQNY 372
QY 362 CDREVEPCIQFSAVQVKNVYENIKGTSATKVAIKEDCSTNPFCEGIIMENILV----- 416
Db 373 CDQKECTQOADAIVNSNMVYQINIRGTSSEVAVKDCSKSPVCGQIILLQDINLVHKGKN 432
QY 417 GESGKPEATCKNVHFNNAEHVTPHC 442
Db 433 DDKSAQAEASCKNVKWKNGRVSPQC 458

RESULT 4
T46187
polygalacturonase (EC 3.2.1.15) precursor [similarity] - Arabidopsis thaliana
N:Alternate names: protein T8H10.110
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 18-Aug-2000
C:Accession: T46187; T50674
R:Benes, V.; Rechmann, S.; Bokkova, D.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23014
A:Accession: T46187
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-431 <BEN>
A:Cross-references: EMBL:AF037367; PIDN:AC98923.1
A:Experimental source: cultivar Landsberg erecta
R:Jenkins, E.S.; Roberts, J.A.
submitted to the EMBL Data Library, December 1997
A:Description: Denescence-related expression of an Arabidopsis thaliana gene encoding a
A:Reference number: Z25172
A:Accession: T50674
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-431 <BEN>
A:Cross-references: EMBL:AF037367; PIDN:AC98923.1
A:Experimental source: cultivar Landsberg erecta
C:Genetics:
A:Gene: T8H10.110
A:Map position: 3
A:Introns: 86/3; 129/3; 186/3; 193/3; 263/1; 290/2; 326/3; 364/3
C:Superfamily: polygalacturonase
C:Keywords: glycosidase; hydrolase
F:1-27/Domain: signal sequence #status predicted <Sig>
F:28-431/Product: polygalacturonase #status predicted <MAT>

Query Match 47.2%; Score 1127; DB 2; Length 431;
Best Local Similarity 49.6%; Pred. No. 2,7e-72;
Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

QY 7 SILLIIIFASSISTCRSNVID-----DNIEFGVYDNILQEFADHQAIVLSYL 55
Db 9 AYFLCVLMLSLCKALSSNVDDGFGHDSFESDSLKANNDVL-----SLISSD 59
QY 56 SKNIESNNNIDKVDKNGIKVINVLISFGAKGDGKTYNDINAFQANWNAACSSRTPEVQFVVPK 115
Db 60 ETLLEAS-----TVSVSNFGAKGDGKTDTPQAFKAMKACSTNGVTLFVPRK 107
QY 116 NKRYLLKQITFSPGCRSSISVTKIFGSLBASKSIDYKDRMLWTFAPDSVQNLVVGG--GT 173
Db 108 GKTYLLKSTYRFSPCKSLRNFQILIGTLSASTKRSIDYDKKNHWLLDENVNLISDGGSTGI 167
QY 174 INNGGVWMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNNV 233
Db 168 INNGKTMWONSCKIKDSKPCCTAFTALITYNLKNLVKYLKRYKNAQOISIEKCNKYE 227
QY 234 ASNLMINASAKSNTDGVHVSNTQYIQISDITIGTDDCISIVSGSQVQATNITCGPH 293
Db 228 VSNVEITAPDPSDNTDGIHTNTQINIRVSNDSIGTDDCISIDGTQNLIDFLTCGPH 287
QY 294 GISISIGSGNSAAYVSNVYVNAKIIIGAENGVRIRIKTMQGGSGQASNIKRLNEMQDYK 353
Db 288 GISISIGSGDNSKAYVSGINVDGAKPSESNDGVRIRIKTYOQSGSTAKNIKQNTIRENVKN 347
```

```
QY 354 PIIDQNYCDREVEPCIQFSAVQVKNVYENIKGTSATKVAIKEDCSTNPFCEGIIMENI 413
Db 348 PIIDQNYCDK-DKEDDQSAVQVKNVYKNTSGTSATDVAITLNCSEKPCQGIIVENV 406
QY 414 NLVGSKPEATCKNVHFNNAEHVTPHCT 443
Db 407 KIKG-----GTASCKRNVKNGVSPKCS 431

RESULT 5
H84846
probable polygalacturonase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: H84846
R:Jin, X.; Kaul, S.; Roumley, S.D.; Shea, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Wolfat, K.S.; Cronin, L.A.; Shen, M.; Yankaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eilen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MID:20083487; PMID:10617197
A:Accession: H84846
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <STO>
A:Cross-references: GB:AE002093; NID:g2335094; PIDN:AA02763.1; GSPDB:GN00139
A:Genes: AE2941850
A:Map position: 2
C:Superfamily: polygalacturonase

Query Match 46.7%; Score 1115.5; DB 2; Length 426;
Best Local Similarity 49.8%; Pred. No. 1,7e-71;
Matches 222; Conservative 74; Mismatches 117; Indels 33; Gaps 7;

QY 6 NSILLIIIFASSISTCRSNVIDNIEFGVY-----DNIEFGVADHQAIVLSYLSKNIE 60
Db 6 NEVTYFLMALMFWSCKASRISPNYDHSYKFKSDSLIKR--EDITGLRSFVNASLR 63
QY 61 SNNNIDKVDKNGIKVINVLISFGAKGDGKTYNDINAFQANWNAACSSRTPEVQFVVPKKNYL 120
Db 64 TPT-----TVSVSDFGAKGDKTDPQAFVNAWKACSSGAVNLVPRKNTYL 112
QY 121 LKQITFSPGCRSSISVTKIFGSLBASKSIDYKDRMLWTFAPDSVQNLVVGG--GTINNG 178
Db 113 LKSIGLTGCONSLTYQIFGTLASQKSDYKDISKIMFEDGVNLSVSGDGVVDGNG 172
QY 179 QVWMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNVVASNL 238
Db 173 EFWQNSCKRKA-----KALTFFNSKSLIVKNLKVNAQOISIEKSNVQVSNV 225
QY 239 INASAKSPNTDGVHVSNTQYIQISDITIGTDDCISIVSGSQVQATNITCGPHGISIG 298
Db 226 VTAPADSPNTDGIHTNTQINIRVSESIIGTDDCISIESGQVQVNDITCGPHGISIG 285
QY 299 SLGSGSEAVSVSNVYVNAKIIIGAENGVRIRIKTMQGGSGQASNIKFLNVEQDYKYPID 358
Db 286 SLGDDNSKAFVSGVYVDGAKLTSTDNGVRIKTYOQSGSGTASNLIFQINQMDVKNPILID 345
QY 359 QNYCDREVEPCIQFSAVQVKNVYENIKGTSATKVAIKEDCSTNPFCEGIIMENILVGE 418
Db 346 QDYCDK-SKCTTEKSAVQVKNVYRDISGSASENAITFCSKNYPCQGIIVLDRVNIKG- 403
QY 419 SGKPEATCKNVHFNNAEHVTPHCTS 444
Db 404 ---GRACTNNAVVDKGAVALPQNS 425

RESULT 6
S71523
polygalacturonase (EC 3.2.1.15) [similarity] - peach
N:Alternate names: endopolygalacturonase
```

C:\Species: Prunus persica (peach)
C:\Date: 19-Mar-1997 #sequence revision 19-Mar-1997 #text_change 31-Mar-2000
C:\Accession: S71523; S71524; S71525
R:\lee, B.; Speirs, J.; Gray, J.; Brady, C.J.
submitted to the EMBL Data Library, April 1994
A:\Description: Homologies to the tomato endopolygalacturonase gene in the peach genome.
A:\Reference number: S71523
A:\Accession: S71523
A:\Molecule type: DNA
A:\Residues: 1-458 <LEW>
A:\Cross-references: EMBL:X77231; NID:S479087; PIDN:CAAS4448.1; PID:S479088
A:\Experimental source: cultivar Maravilla
A:\Note: this is a revision to the sequence from reference S71524
R:\lester, D.R.; Speirs, J.; Orr, G.; Brady, C.J.
submitted to the EMBL Data Library, January 1994
A:\Description: Homologies to the tomato endopolygalacturonase gene in the peach genome.
A:\Reference number: S71524
A:\Accession: S71524
A:\Molecule type: DNA
A:\Residues: 53-92; 'I', '94-149', 'T', '151-185', 'W', '187-196', 'THA', '200', 'ESLS', '206-207', 'TNT', '212
A:\Cross-references: EMBL:X77231
A:\Experimental source: cultivar Maravilla
A:\Note: this sequence has been revised in reference S71523
R:\lester, D.R.; Speirs, J.; Orr, G.; Brady, C.J.
Plant Cell Env. 13, 513-521, 1990
A:\Title: Homologies to the tomato endopolygalacturonase gene in the peach genome.
A:\Reference number: S71525
A:\Accession: S71525
A:\Molecule type: DNA
A:\Residues: 265-458 <LEW>
A:\Cross-references: EMBL:X77231
A:\Experimental source: cultivar Maravilla
C:\Genetics:
A:\Intons: 107/3; 150/3; 204/3; 213/3; 283/1; 310/2; 386/3
C:\Superfamily: polygalacturonase
C:\Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 46.4%; Score 1108; DB 2; Length 458;
Best Local Similarity 48.9%; Pred. No. 6.5e-71;
Matches 226; Conservative 82; Mismatches 126; Indels 28; Gaps 8;

OY 1 MTIGRNSILLLIITIPASISTCRSNVIDNLFKOYDNIIEOFADHFOAYLSYLS---- 56
DB 1 MALOKHVLFLVYVVSFCFCAACYSGSGFOEVNLSLHFSVDH--EKESGYNSRAHPNMNTEG 58
OY 57 -----KNISNNNDKV---DKNGIVINVLSSGGAQGDGKTYDNIATFEQANMEAC 103
DB 59 VKMEFIKPRAPQLSSRKLERAGSKSSSVTTISVANFGAGANGAD--DTRAFEXAKMAAC 117
OY 104 SSRPVOFVPPKKNKYNLYLKQITTFGPPCRSSISVYKIFGSLSEAKSIDYKDKRLWIAFPSV 163
DB 118 SSNGAIVLVLVPPQ-KTIVLRPIEFSGPCKSHLTMOYIGTIESSDDRSYKVDVTHMLIHPNV 176
OY 164 QNLVVGSGGTTNGNGOVWVPSSCKINKSLPCRD-APFALITFNCNKILKVNLIKSKNAQOI 222
DB 177 QSLIVVPPGTINGNGRWMWENSCKRKPQPCNEQAPTAFTNKKCNILVLVNLKIQDAQOM 236
OY 223 HIKESCTNVVASTLMTNLSAKSPNTDGVHNSNTQYIQISPTIIGTGGDCTSVSGSNV 282
DB 237 HVRFQNCNVEASHLVTAPEDSPNTDGIHITNKNKITISSVIGTGGDCTSVSGSORV 296
OY 283 QATNITGCPGHGISIGSLGSGNSEAVYSNVTVNEAKTIGAENGVRKITWGGSGOASNIK 342
DB 297 QATDITGCPGHGISIGSLGEDNANDHVSIGVFNAGAKISGTSGNGRITKWTWGGSGSASNV 356
OY 343 FLAVEMODVKKPIITIDNYCD-RVEPCIQPSAYQVKNVYENIKGTSATVAIKFDOST 401
DB 357 FQVEMNDVNPPIITIDONYCDHKKDKDCTGRSAVQVKNVLYQNTRGTSASDATTFCOSQ 416
OY 402 NPFCEGLIMENINLVGSGKPEATCKNVHNNNAEHTPHCT 443
DB 417 SVPPQGIIVDNLQ-----QVARAKCNVAPRAYGAVSPPRS 453

RESULT 7
T17011
polygalacturonase (EC 3.2.1.15) - apple tree
N|Alternate names: poly [1,4-a-D-galacturonide] glycan hydrolase
C|Species: Malus domestica (apple tree)
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C|Accession: T17011
R|Atkinson, R.G.
Plant Physiol. 105, 1437-1438, 1994
A|Title: A cDNA clone for endopolygalacturonase from apple.
A|Reference number: Z18649; MUID:95062722; PMID:7972500
A|Accession: T17011
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: mRNA
A|Residues: 1460 <Atk>
A|Cross-references: EMBL:L27743; NID:g456091; PIDD:AAA74452.1; PID:g456092
A|Experimental source: Strain Golden delicious; ripe fruit
C|Function:
A|Description: catalyzes hydrolysis of 1,4-alpha-D-galactosiduronic linkages in galactur
A|Pathway: polysaccharide degradation
C|Superfamily: polygalacturonase
C|Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 45.6%; Score 1091; DB 2; Length 460;
Best Local Similarity 49.3%; Pred. No. 1,1e-69;
Matches 231; Conservative 76; Mismatches 126; Indels 36; Gaps 12;

QY 1 MWQRNSILLILITPAASIST--CRSNVTD-NLFQVYDNLIEQFAHDFQAYLSLK 57
DB 1 MALKTQLMSFVVEVVSFSTSCSSSFQEVVALHSYVDHVDKSGVSRAPSY-TD 59
QY 58 NIE-----SNNIDKVDKNGI-----VIVNLSGAKGDKGTYNIAF 95
DB 60 TTEGKTMELLRPRTQLFPSSRKLTNT-TGCLITSSAPAKTISVDDPGAKNGAD-DTQAF 117
QY 96 EQAWNACSSRTPVQVFPVKKNKYLKQITFSGPCSSISVKLFGSLBASISKIDYKDR 155
DB 118 VKAWKPAACSSSGAMVLVYPQ-KNYLVRIEFGPCSGQLTLQVGTLEASGDSIYDID 176
QY 156 LMIAPDSYQNLVVGSGGTTNGQVWPBSSCKINSLPRD-APTALTFFNCKKLYNNL 214
DB 177 HMLTFDNYQNLVVGPGTNGNGNIMWKNKSKIKPQPCGTVAFLVTFNRKNLVLVKNL 236
QY 215 KSKNAQQLHIKPESSNTNVASNLMINASAKSNTDGVHSNQCYSIDRTIIGTDDCIS 274
DB 237 NIQDQQLHVFQNCINYOASCLITVTAPEPDSNPDGSHVNTONTITSSVIGTDDCIS 296
QY 275 IVSGSQNVQATNTTCGPHGISIGSLGSGNSEAVYSNVVNEAKITGAENGVRIKTWQG 334
DB 297 IVSGSQRVQADTDITCGPHGISIGSLGSGSDSHVSGVFNKAKSGTSGLRKIKTWKG 356
QY 335 SGOASNIKFLANEMQDVKPIIIDQNYCD-KVPEPCIQGSAYQKVVVENIIGTSATKV 393
DB 357 SGGATNIVQVNMVDVNPPIIIDQNYCDHKTKDKQQAQVKNVLYONIGTSASGD 416
QY 394 AIFKDCNPFCEGIGIMENINLVGSGKPEATCKVHNNAEHWPHG 442
DB 417 AITLNGSQSVCEQGILOSVDL--QNGR---AFCNVQVPAKGVSPRC 460

RESULT 8
D96833
hypochemical protein P1813.25 [imported] - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C|Accession: D96833
R|Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chai, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A|Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; A.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzitali,

107 VERGPCOHLSEKVDGTTVAQDPAPRWKSKIMLOFAQLTDFNLMTGVIDGQGGQWMAAG 166
QY 185 SCK-INKSLPC--RDAPALTFFWCKNKLKVNLSKNAQOIHKEESCTNVASNLMTNA 241
Db 167 QCKVWNGSTVCNDRRPRPAIKIDYKSVYKELTLMNSPEFLVGECEGYKIQGLKTKA 226
QY 242 SAKSPNTDGVHVSNTQYIQTSDTIIGTDDCISIVSGSQNVQATNITGCPHGIGISGLG 301
Db 227 PRDSPNTDGDIFASKRFHIEKVCVIGTDDCIALIGTSSNITIKDLIGCPHGIGISGLG 286
QY 302 SGNEEAYVSNVTWNEAKTIGENGVRIKTKWGGSGQASNIKFLNVEWQVXPIIIDONY 361
Db 287 RDNRAEYSHVHVNAKFIIDTQNGRLIKTKWGGSGLASIYVENEMTNSNPILINOFY 346
QY 362 CDRVEPCIQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESGK 421
Db 347 CTSASACQNGRSAVQIDQVYTKNIGHTSATAAAIQLMCSDSVPCIGTQLSNVSLKLTSGK 406
QY 422 PSEATCKNVHNNNAEHVTPHCTSL 445
Db 407 PASCVDKXNARGFYSGRLLIPTCKNL 430

RESULT 11

A96609
probable polygalacturonase F25P12.85 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C/Accession: A96609
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huntz, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: A96609
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-434 <STO>
A/Cross-references: GB:AE051173; NID:g9954742; PID:AA09093.1; GSPDB:GN00141
C/Genetics:
A/Gene: F25P12.85
A/Map position: 1
C/Superfamily: polygalacturonase

Query Match 32.9%; Score 787.5; DB 2; Length 434;
Best Local Similarity 41.3%; Pred. No. 3.1e-48;
Matches 170; Conservative 68; Mismatches 155; Indels 19; Gaps 8;

QY 54 YLSKNITSNNTIDKVDKNGIKVINLSPGAKGDKTYDNIAPFEOAMNEACSSRTP--VQF 111
Db 26 YLSPSPAPNPAVYNDNDIAPVFDVTSFGAIGDCSTDTSAFKAMWDAACMSTSPKALL 85
QY 112 VVPEKNKYLLKQITFSGPCRSISVYKIFGSLA-----SSKIDVKDRRLMIAFVSQNTLV 167
Db 86 LVPTTFCFLKPTTFNPCTNVLQIDGFIIVSDGPRSPSNY--QRWMMFRRVNGLS 143
QY 168 VGGGCTINGNGQVWVWSPSSCKINKSL-----PCRDAPALTFFWCKNKLKVNLSKNAQ 220
Db 144 IQSGGVINGGQKWMNLPCPKHKLNGTQTGFC-DSFVALIRLFGSSKVRIGGINPMSA 202
QY 221 QIHKEESCTNVASNLMTNAKSPNTDGVHVSNTQYIQTSDTIIGTDDCISIVSGSQ 280
Db 203 QFHVRFNCDVVDVSVIIKAPASSPMTDGIHENTHNVQIRNMSISNGDCISIGAGCF 262
QY 281 NVQATNITGCPHGIGISIGSGNSEAYVSNVTWNEAKTIGENGVRIKTKWGGSGQASN 340
Db 263 NVDIKNTVCGPSHGIGISIGSGVHNSQAVSNITVTNSTIWNDSGVRIKTKWGGSGSVSR 322

QY 341 IKELNVEWQVXPIIIDONYCDRVEPCIQFSAVQVKNVYENIKGT-SATKVAIKFDC 399
Db 323 IVFSNITLVNVRNPMIDQYIC-QTNNCANQTSATVISTIDYANKITGLDRSPPIHFGC 381
QY 400 STNFPCEGIMENINLVGESGKPSF-ATCKNVHFNNAHVTPHCTSLISED 450
Db 382 SDSVPCNTLTLTEVDLFPSSKQGHLENPFCWNAVSGMKITVPPVYCLLAPPD 433

RESULT 12

S48730
Cry j II protein - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C/Accession: S48730
R/Nanda, M.; Kurose, M.; Torigoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kur
FEBS Lett. 353, 124-128, 1994
A/Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar p
A/Reference number: S48730; MUID:95010777; PMID:7926035
A/Accession: S48730
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-514 <NMA>
A/Cross-references: GB:D37765; NID:g577695; PID:BA07021.1; PID:d1007598; PID:g577696

Query Match 32.9%; Score 787.5; DB 2; Length 514;
Best Local Similarity 38.5%; Pred. No. 3.9e-48;
Matches 173; Conservative 78; Mismatches 163; Indels 35; Gaps 9;

QY 8 ILIIIIIFASSISYCRSNVIDNLFEQYVDNILEQEFNAHDFQAVLYSKNIESNNNDK 67
Db 12 VAMQILIMAA-----EDQSAQIMLSDIDQ-----YL-----RNSRSIRK 47
QY 68 VD---KNGIKVINLSPGAKGDKTYDNIAPFEOAMNEACSSRTPVQVVPKKNYLLKQI 124
Db 48 VEHSHDAININFEVEKYGAVDGDGHDCTEAFSTAWQAAC-KKPSAMLVPPGNKKFVVNVL 106
QY 125 TFSGCRSSISVYKIFGSLSEASSKISDYKDRRLMIAFVSQNTLVVGGGCTINGNGQVWVPS 184
Db 107 FTFNPGCPHFHFFKVDGILIAAQNPASWKNRIMVQFAKLTGFTLMGKVIDIGQGGQWMAAG 166
QY 185 SCK-INKSLPC--RDAPALTFFWCKNKLKVNLSKNAQOIHKEESCTNVASNLMTNA 241
Db 167 QCKVWNGREICNDRRPRPAIKFDPSTGILLIGKLTMSPEFLVGECEGYKIIIGISTTA 226
QY 242 SAKSPNTDGVHVSNTQYIQTSDTIIGTDDCISIVSGSQNVQATNITGCPHGIGISGLG 301
Db 227 PRDSPNTDGDIFASKRFHIEKVCVIGTDDCIVALGTSSNITVIDLIGCPHGIGISGLG 286
QY 302 SGNEEAYVSNVTWNEAKTIGENGVRIKTKWGGSGQASNIKFLNVEWQVXPIIIDONY 361
Db 287 RDNRAEYSHVHVNAKFIIDTQNGRLIKTKWGGSGMASHIYVENEMTNSNPILINOFY 346
QY 362 CDRVEPCIQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESGK 421
Db 347 CTSASACQNGRSAVQIDQVYTKNIGHTSATAAAIQLMCSDSVPCIGTQLSNVSLKLTSGK 406
QY 422 PSEATCKNVHNN--AEHVTPHCTSL 448
Db 407 --IASCLNDNANGVYSGHVPACNKLSPS 433

RESULT 13

JC2498
second major allergen Cry j II precursor - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C/Accession: JC2498; PC346; A60147
R/Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A/Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese
A/Reference number: JC2498; MUID:94271186; PMID:8002972

A:Accession: J02498
A:Molecule type: mRNA
A:Residues: 1-514 <KOM>
A:Cross-references: DDBJ:D29772; NID:g506857; PIDN:BAA06172.1; PID:g506858
A:Accession: PC3346
A:Molecule type: protein
A:Residues: 52-61 <KO2>
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matubasi, T.
Allergy 45, 309-312, 1990
A:Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147; MUID:90342988; PMID:2382797
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAX>
C:Keywords: glycoprotein; pollen
F:1-54/Domain: signal sequence #status predicted <Sig>
F:55-60/Product: second major allergen Cry j #status predicted <MAT>
F:429,460/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 32.9%; Score 787; DB 2; Length 514;

Best Local Similarity 40.5%; Pred. No. 4,2e-48;

Matches 168; Conservative 73; Mismatches 160; Indels 14; Gaps 7;

```
QY 45 AHDPQAYL---SYLSKNISSNNIDKYD---KNGIVYIVLSPGAGDQKTYDNIAPEQ 98
DB 22 AEDQSAQIMDSVEKYLRSNLSRKHESRHDALNIFVWEKYGAGDGHDCTEAFSTA 81
QY 99 WNEACGSRTPVQVYVKNKNVYLKQITFSGPCSSISVKIFGSLBASKISDYKDRRLMI 158
DB 82 WQACACNPS-AMLLVGSKKFVNNLFPNGPCOPHFTFYKVDGILAYYONPASKNNRLMI 140
QY 159 AFDVQNLVVGGGGTNGNGQVWPPSSCK-INKSLPC--RDAPTALTFFNCKNLKYNLXK 215
DB 141 QFAKLITGFTLMGKGVLDGGCKQWAGQCKVWNGRETCNDRRTALKFDPSTGLIIGXK 200
QY 216 SKNAQOIHFKFESCTNVASNMNINSAKSPNTDGVHVSNTQYIQISDTIIGTDDICSI 275
DB 201 INMSPEFHVFGNCECVKIIIGISITRAPDSPNDGIDIFASKNPHLOKMTIGTDDCVAI 260
QY 276 VSSGQVQVQNTITCGPHGISIGSLGSGNSEAVSVTVTEAKITGAENVRIRKTMQGS 335
DB 261 GTSSSVIVIEDLTCGPHGISISLRNSRAVSVHVNGAFPIDTQNGRLKTMQGS 320
QY 336 GQASNIKFLNVEQVQYPIIDQNYCDREPCIOQPSAVQVKNVYENIKGTSATKVAI 395
DB 321 GMAASHIYENINENPILINQFYCTASACQNSANQIDQVYKNIKRGTSATPAAI 380
QY 396 KPDGCTNFPCEGIMENINLVGSGKPSKSEATCKNVHFN--AEHVTPHCTSLKIS 448
DB 381 QKCSDSMPCKDKLSDISIKLTSK--IASCLINDNANGVFGSHVLPACKNLSPS 433
```

RESULT 14

JC7366

Jun a 2 protein - mountain cedar

C:Species: Juniperus ashei (mountain cedar)

C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C:Accession: JC7366; PC7093

R:Tokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A:Title: Purification, identification, and cDNA cloning of Jun a 2, the second major a11

A:Reference number: JC7366

A:Accession: JC7366

A:Molecule type: mRNA

A:Residues: 1-507 <YOK>

A:Cross-references: GB:AJ0404653

A:Accession: PC7093

A:Molecule type: protein

A:Residues: 55-63 <Y02>

C:Comment: This protein, a second major allergen of mountain cedar pollen, which is inv

C:Keywords: glycoprotein; pollen

Query Match 32.0%; Score 764; DB 2; Length 507;

Best Local Similarity 41.8%; Pred. No. 1.8e-46;

Matches 157; Conservative 66; Mismatches 145; Indels 8; Gaps 5;

```
QY 75 VINVLSFGAKGDKTYDNIAPQAWNEACSSRTPVQVYVKNKNVYLKQITFSGPCRSSI 134
DB 59 VNEVHYGAVGDKDSTDAFEKTMMAACNKL SAV-FLVPANKKVVNNLVFPGCPQPH 117
QY 135 SVKIFGSLBASKISDYKDRRLMIAFDVQNLVVGGGGTNGNGQVWPPSSCK-INKSLP 193
DB 118 SFKVDGTLIAYPDKPAKXSKIMHFPARLTDNLMTGIVLDGGKRWMSDQCKTINGRV 177
QY 194 GCD--APYALTFFNCKNLKYNLXKNAQOIHFKFESCTNVASNMNINSAKSPNTDGV 251
DB 178 CNDKGPPTAIKIDFSKSVTKELTLNLSPEHFLVFECCGVKIQGKIKAPRSPYTDGI 237
QY 252 HYSNTQYIOISDTIIGTDDCISYVSGSQVQATNITCGPHGISIGSLGSGNSEAVSVN 311
DB 238 DIFASRFEIEKCTIGTGDVAVGSGSNITIKDLTCGPHGMSIGSLGSKNSRSVSR 297
QY 312 VTVNEAKITGAENGVARIKTWQSGQASNIKFLNVEQVQYPIIDQNYCDREPCIOQ 371
DB 298 VELDQAKFTDTQNGRLKTMQGSGLASHIYENINENPILINQFYCTASAAQXQ 357
QY 372 FSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGSGKPSKSEATCKVH 431
DB 358 RSAVKIQDVTFFKNHGTSAITTAIQLMCSDSVCSNKLKSLNFKLTSK--VATCVNKN 415
QY 432 ENN--AEHVTPHCTSL 445
DB 416 ANGYTTPNLNPSCKSL 431
```

RESULT 15

B86368

protein F28C11.9 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B86368

R:Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltl, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Frazer, C.W.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B86368

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1161 <STO>

A:Cross-references: GB:AE005172; NID:g8778576; PIDN:AAF79584.1; GSPDB:GN00141

A:Gene: F28C11.9

A:Map position: 1

Query Match

31.4%; Score 749.5; DB 2; Length 1161;

Best Local Similarity 36.1%; Pred. No. 5.8e-45;

Matches 163; Conservative 73; Mismatches 133; Indels 83; Gaps 8;

```
QY 71 NGIKVINVLSFGAKGDKTYDNIAPQAWNEACSSRTPVQVYVKNKNVYLKQITFSGPC 130
DB 708 SGNLVNVDPTFGAAGGVSDDTQAFVSAMSKASTSKSV-FLVPEGRRIYVNAITKFGPC 766
QY 131 RSSISV-----KIFGSLBASKIS--DYKDRRLW 157
DB 767 EQLTIIQVKLSSVGKQXLDHNYIYKYVLVNSFTCLQIDGTIVAPEPNSMWSKFGRIW 826
QY 158 IAFDSVQNLVVGGGGTNGNGQVWPPSSCKINSLPORDAPYALTFFNCKNLKYNLXSK 217
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:02:02 ; Search time 15.6574 Seconds
(without alignments)
1519.798 Million cell updates/sec

Title: US-10-018-604-2

Perfect score: 2390

Sequence: 1 MWIQRSILLIIIFASSIS.....VTPHCTSLIISDEALVNY 457

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2390	100.0	457	1 PGLR_LYCES	P05117 lycopersico
2	1295.5	54.2	467	1 PGLR_ACTCH	P35336 actinidia c
3	1136	47.5	462	1 PGLR_PEAR	Q02096 persea amer
4	1091	45.6	460	1 PGLR_MALDO	P48978 malus domes
5	787.5	32.9	514	1 MPAA_CRYUA	P43212 cryptomeria
6	708.5	29.6	422	1 PGL1_ARATH	P49062 arabidopsis
7	687.5	28.8	393	1 PGLR_PRUPE	P48979 prunus pers
8	667	27.9	447	1 PGLR_GOSHI	Q39786 gossypium h
9	667	27.9	444	1 PGL2_ARATH	P49063 arabidopsis
10	666	27.9	407	1 PGLR_GOSBA	Q39766 gossypium b
11	650.5	27.2	410	1 PGLT_MAIZE	P35339 zea mays (m
12	630	26.4	362	1 PGLR_ORNOR	P24548 ornithera o
13	627	26.2	410	1 PGLR_MAIZE	P26215 zea mays (m
14	626	26.2	410	1 PGLS_MAIZE	P35338 zea mays (m
15	589.5	24.7	421	1 PGLR_MEDSA	Q40312 medicago sa
16	582.5	24.4	396	1 PGLR_TOBAC	Q05967 nicotiana t
17	572.5	24.0	397	1 PGLR_BRANA	P35337 brassica na
18	346	14.5	446	1 PGLX_COCCA	Q00339 cochlioboln
19	313	13.1	353	1 PGLR_ASFOR	P35335 aspergillus
20	311	13.0	370	1 PGL1_PENOL	Q9Y834 penicillium
21	311	13.0	435	1 PGLX_ASPTU	Q00293 aspergillus
22	310	13.0	363	1 PGLA_ASPTU	P41749 aspergillus
23	306	12.8	363	1 PGLR_ASPTU	P49575 aspergillus
24	300	12.6	364	1 PGLR_COCCA	P26215 cochlioboln
25	300	12.6	367	1 PGLR_PENDI	Q9Y718 penicillium
26	296	12.4	361	1 PGLR_YEAST	P47180 saccharomyc
27	291.5	12.2	402	1 PGL2_ERWCA	P26509 erwincia car
28	287.5	12.0	402	1 PGL1_ERWCA	P18192 erwincia car
29	285.5	11.9	368	1 PGL1_ASPTU	P26213 aspergillus
30	285	11.9	362	1 PGL2_ASPTU	P19805 aspergillus
31	280	11.7	376	1 PGLR_PENGR	Q93883 penicillium
32	278	11.6	363	1 PGL1_COLAN	Q00446 colletotric
33	277	11.6	380	1 PGL1_SCLIS	Q12708 sclerotinia

34	275	11.5	366	1 PGLB_ASPTU	P41750 aspergillus
35	273.5	11.4	371	1 PGLR_PENNA	Q42824 penicillium
36	271.5	11.4	373	1 PGLR_GIBFU	Q07181 gibberella
37	265	11.1	362	1 PGL2_ASPTU	P26214 aspergillus
38	256	10.7	380	1 PGL2_PENOL	Q9Y833 penicillium
39	247	10.3	383	1 PGL3_ASPTU	Q12554 aspergillus
40	242.5	10.1	378	1 PGLR_PENEN	Q59825 penicillium
41	241	10.1	440	1 RHGA_ASPTU	Q00001 aspergillus
42	237.5	9.9	602	1 PEHX_ERWCA	P15922 erwincia chr
43	186.5	7.8	531	1 PGL1_RALSO	P58598 ralstonia s
44	183	7.7	312	1 PGLR_AGRUT	P27644 agrobacteri
45	181	7.6	529	1 PGL2_RALSO	P20041 ralstonia s

ALIGNMENTS

RESULT 1

PGLR_LYCES STANDARD; PRT; 457 AA.

AC P05117;
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Polygalacturonase 2A precursor (EC 3.2.1.15) (PG-2A) (pectinase).
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Ailsa Craig;
 RA Bird C.R., Smith C.J.S., Ray J.A., Mourreau P., Bevan M.W.,
 RA Bird A.S., Hughes S., Morris P.C., Grierson D., Schuch W.;
 RT "The tomato polygalacturonase gene and ripening-specific expression in
 RT transgenic plants.";
 RL Plant Mol. Biol. 11:651-662(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Ailsa Craig;
 RX MEDLINE=8706731; PubMed=3786135;
 RA Grierson D., Tucker G.A., Keen J., Ray J., Bird C.R., Schuch W.;
 RT "Sequencing and identification of a cDNA clone for tomato
 RT polygalacturonase.";
 RL Nucleic Acids Res. 14:8595-8603(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Sheehy R.E., Pearson J., Brady C.J., Hiatt W.R.;
 RT "Molecular characterization of tomato fruit polygalacturonase.";
 RL Mol. Gen. Genet. 208:30-36(1987).
 RN [4]
 RP REVISIONS.
 RA Hiatt W.R.;
 RT Submitted (OCT-1987) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Acts in concert with the pectinesterase, in the ripening
 CC process. Is involved in cell wall metabolism, specifically in
 CC polyuronide degradation.
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
 CC galactosiduronic linkages in pectate and other galacturonans.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DEVELOPMENTAL STAGE: In ripening fruit.
 CC -1- BIOTECHNOLOGY: The effect of PG can be neutralized by introducing
 CC an antisense PG gene by genetic manipulation. The Flavr Savr
 CC tomato produced by Calgene (Monsanto) in such a manner has a
 CC longer shelf life due to delayed ripening.
 CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
or send an email to license@isb-sib.ch).

DR EMBL, X14074; CAA32235.1; -
DR EMBL, M37304; AAA34178.1; -
DR EMBL, X04583; CAA28254.1; -
DR EMBL, X05656; CAA29148.1; -
DR EMBL, A15981; CAA01256.1; -
DR EMBL, A24194; CAA0120.1; -
DR PIR, A25534; A25534.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; PDH1.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SMO0710; Pdh1; 4.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
KW Hydrolyase, Glycosidase; Cell wall; Signal; Fruit ripening;
KW Glycoprotein; Genetically modified food.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 71 POLYGALACTURONASE 2A.
FT CHAIN 72 457
FT ACT SITE 293 293 PROBABLE. (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 457 AA; 50051 MW; 449EADCC6919B074 CRC64;

Query Match 100.0%; Score 2390; DB 1; Length 457;
Best Local Similarity 100.0%; Pred. No. 1,6e-160;
Matches 457; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1 MVIORNSILLIIIFASSISIFCRSNVIDNLFKQYVDNLTBEFPAHDFQAYLSLTKNTE 60
DB 1 MVIORNSILLIIIFASSISIFCRSNVIDNLFKQYVDNLTBEFPAHDFQAYLSLTKNTE 60

QY 61 SNNNIDKVDKNGIKVINLVSFGAKDGKTYDNIAEQAMNACSSRTPEVGFVPEKKNYLT 120
DB 61 SNNNIDKVDKNGIKVINLVSFGAKDGKTYDNIAEQAMNACSSRTPEVGFVPEKKNYLT 120

QY 121 LKQTFPGPCSSISVXIFGSLSEASKISDYKDRRLMTAFPSVONLVVGGGGINNGQV 180
DB 121 LKQTFPGPCSSISVXIFGSLSEASKISDYKDRRLMTAFPSVONLVVGGGGINNGQV 180

QY 121 LKQTFPGPCSSISVXIFGSLSEASKISDYKDRRLMTAFPSVONLVVGGGGINNGQV 180
DB 121 LKQTFPGPCSSISVXIFGSLSEASKISDYKDRRLMTAFPSVONLVVGGGGINNGQV 180

QY 181 WMPSSCKINKSLPRDAPALTLPNNCKNLKYNLKSXAQOIHFKFSCCTVVASNLMIN 240
DB 181 WMPSSCKINKSLPRDAPALTLPNNCKNLKYNLKSXAQOIHFKFSCCTVVASNLMIN 240

QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTGDCCISIVSGSQNVQATNITCGFGHGISIGSL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTGDCCISIVSGSQNVQATNITCGFGHGISIGSL 300

QY 301 GSGNSEAVYVNTVNEAKITGAENGVRKTMQGGSGQASNIKFLNVEMQDYKYPIIIDON 360
DB 301 GSGNSEAVYVNTVNEAKITGAENGVRKTMQGGSGQASNIKFLNVEMQDYKYPIIIDON 360

QY 361 YCDREVEICIOQFSAVQKNNVYENIKGTSATKVAIKFDCSTNFCDEGIMENINLVGESG 420
DB 361 YCDREVEICIOQFSAVQKNNVYENIKGTSATKVAIKFDCSTNFCDEGIMENINLVGESG 420

QY 421 KPSEATCKNVFNNAEHVTPHCTSLSEIDEDALLNNY 457
DB 421 KPSEATCKNVFNNAEHVTPHCTSLSEIDEDALLNNY 457

RESULT 2
PGLR ACTCH STANDARD; PRT; 467 AA.
AC P35336;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polyalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase).
OS Actinidia chinensis (Kiwi) (Yangtze).

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
CC Ericales; Actinidiaceae; Actinidia.
OK NCBI_TaxID=3625;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Deliciosa;
RX MEDLINE=94302157; PubMed=8029342;
RA Atkinson R.G., Gardner R.C.;
RT "A polyalacturonase gene from kiwifruit (Actinidia deliciosa).";
RL Plant Physiol. 103:669-670(1993).
CC -!- FUNCTION: Acts in concert with the pectinesterase, in the ripening
CC process. Is involved in cell wall metabolism, specifically in
CC polyuronide degradation.
CC -!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: In ripening fruit.
CC -!- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
or send an email to license@isb-sib.ch).

DR EMBL, L12019; AAC14453.1; -
DR InterPro; IPR000743; Glyco_hydro_28.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
KW Hydrolyase, Glycosidase; Cell wall; Signal; Fruit ripening;
KW Glycoprotein.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 467 POLYGALACTURONASE.
FT ACT SITE 306 306 PROBABLE. (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 467 AA; 50776 MW; 5A9A61483C028B7A CRC64;

Query Match 54.2%; Score 1295.5; DB 1; Length 467;
Best Local Similarity 54.6%; Pred. No. 1,1e-83;
Matches 257; Conservative 75; Mismatches 112; Indels 27; Gaps 7;

QY 1 MVIORNSILLIIIFASSISIFCRSNVIDNLFKQYVDNLTBEFPAHDFQAYLSLTKNTE 58
DB 1 MVIORNSILLIIIFASSISIFCRSNVIDNLFKQYVDNLTBEFPAHDFQAYLSLTKNTE 58

QY 59 IESNNIDKVDKNGI-----KVYNLVSFGAKDGKTYDNIAEQAMN 101
DB 59 IESNNIDKVDKNGI-----KVYNLVSFGAKDGKTYDNIAEQAMN 101

QY 102 ACSSRTPEVGFVPEKKNYLLKQTFPGPCSSISVXIFGSLSEASKISDY-KDRRLMTAF 160
DB 102 ACSSRTPEVGFVPEKKNYLLKQTFPGPCSSISVXIFGSLSEASKISDY-KDRRLMTAF 160

QY 115 ACSSTSAVLLVPR-KVYLVRPIFSFGCKSGLTQMTQYGTIEASDDSDYKDRHMLVF 173
DB 115 ACSSTSAVLLVPR-KVYLVRPIFSFGCKSGLTQMTQYGTIEASDDSDYKDRHMLVF 173

QY 161 DSVONLVVGGGGINNGQVWMPSSCKINKSLPRDAPALTLPNNCKNLKYNLKSXAQ 220
DB 161 DSVONLVVGGGGINNGQVWMPSSCKINKSLPRDAPALTLPNNCKNLKYNLKSXAQ 220

QY 174 DSVONLVVGGGGINNGQVWMPSSCKINKSLPRDAPALTLPNNCKNLKYNLKSXAQ 220
DB 174 DSVONLVVGGGGINNGQVWMPSSCKINKSLPRDAPALTLPNNCKNLKYNLKSXAQ 220

QY 221 QIHIFESCTNVVASNLMINASAKSPNTDGVHVSNTQYIQISDTIIGTGDCCISIVSGSQ 280
DB 221 QIHIFESCTNVVASNLMINASAKSPNTDGVHVSNTQYIQISDTIIGTGDCCISIVSGSQ 280

QY 234 QIHVSFDCNVVQASNLMTAFAPENSPTDGIHVTGQNHHSVCVIGTGDCCISIVSGSR 293
DB 234 QIHVSFDCNVVQASNLMTAFAPENSPTDGIHVTGQNHHSVCVIGTGDCCISIVSGSR 293

QY 281 NVQATNITCGFGHGISIGSLGSGNSEAVYVNTVNEAKITGAENGVRKTMQGGSGQASN 340
DB 281 NVQATNITCGFGHGISIGSLGSGNSEAVYVNTVNEAKITGAENGVRKTMQGGSGQASN 340

QY 294 KVRVNDITCGFGHGISIGSLGSGNSEAVYVNTVNEAKITGAENGVRKTMQGGSGQASN 353
DB 294 KVRVNDITCGFGHGISIGSLGSGNSEAVYVNTVNEAKITGAENGVRKTMQGGSGQASN 353

QY 341 IKFLNVEMQDYKYPIIIDONVCDREVEICIOQFSAVQKNNVYENIKGTSATKVAIKFPCS 400
DB 341 IKFLNVEMQDYKYPIIIDONVCDREVEICIOQFSAVQKNNVYENIKGTSATKVAIKFPCS 400

QY 354 IKFQNVEMHVENVTIIDONVCDODKPCQDQSSAVQKNNVYENIKGTSATKVAIKFPCS 413
DB 354 IKFQNVEMHVENVTIIDONVCDODKPCQDQSSAVQKNNVYENIKGTSATKVAIKFPCS 413

QY 401 TNFCDEGIMENINLVGESGKPSEATCKNVFNNAEHVTPHCTSLSEIDE 451

[illegible]

Qy	28	DNULFKQVYNNIIIEGFADHFAVLYSL,SKNIESNNNDIKVNRGCIKYNVLSFAKGG	87
Db	40	DQSAPIYFEPILID-EFSSIMGFEBILS--LERNFVGGEPETSPDTLISVDDFGARGG	96
Qy	88	KTYDNIAFEQAWNEACSSRTPVQFVVPKNKNYLLKQITFGPCRSSISYKIFGSLBASCK	14
Db	97	-TDDTAPEAKMDACSSGVL--LYPEKNYLLKQITFGPCKSDLRYKINGITBASD	15
Qy	148	ISDY--KDRLMIAFDSVQNLVYGGGGITNGNQVWMBSSCKINSLPERDAPITALTFMN	20
Db	154	QSDWVGNNRRKRMWIEPEDISNLTLEGGSITNGNGETWMBSSCKRKSLLPEKSAPTALTFRS	21
Qy	206	CKULXVNNLSKNAQOQIHKFESCTVWVASNNIMNASKSPYTDGVAHVNTQYIOISDPI	26
Db	214	CKULIYSDLSIKOSQCKHLSPDCQVIVASNLWVLAPEHSYPTDGIHTTGTRIHVANSV	27
Qy	266	IGTGDPCISIVSGSQNVQATNITCGPHGHSIGSLSGNSEAVYVSNVTVNEAKIIGAENG	32
Db	274	IGTGDDCISIESSSKXVIALNITCGPHGHGISIGSLDRNSEAHVSGSVLVDGGLPDTTNG	33
Qy	326	VRIKTWQSSGQASNIKFLNVEMQVYPIIIDQNYCDREVPCTIQQFSAVQKRVYENI	38
Db	334	LRKLTWQSSGGSAMNIKEQNIWVHNTNPIIIDQYCDSEKDCPEQESAVKVSNAVMNI	39
Qy	386	KGTSAATVAIKPDCSTNPECEGIMENINUNESGKPSATCKNV	430
Db	394	RGTSAEVAVKPDCSKSSPCQGTIVANINLVNGGKELTMSGNI	438

```

RESULT 4
PGLR_MALDO
ID _PGLR_MALDO STANDARD; PRF; 460 AA.
AC P46978;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Polysialacturonase precursor (EC 3.2.1.15) (PG) (Pectinase) .
OS Malus domestica (Apple) (Malus sylvestris)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosids I; Rosales; Rosaceae; Maloideae; Malus.
OC
OX NCBI_TaxID=3750;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Golden Delicious;
RX MEDLINE=95062722; PubMed=7972500;
RA Atkinson R.G.;
RT "A cdna clone for endopolysialacturonase from apple.";
RL Plant Physiol. 105:1437-1438(1994) .
CC -I- FUNCTION: acts in concert with the pectinesterase, in the ripening
CC process. Is involved in cell wall metabolism, specifically in
CC polyuronide degradation.
CC -I- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L27743; AAA74452.1; -.
DR PIR; T17011; T17011.
DR InterPro; IPR000743; Glyco_hydro_28.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR PROSITE; PS00502; POLYSIALACTURONASE; 1.
DR Hydrolase; Glycosidase; Cell wall; Signal; Fruit ripening;
KM Glycoprotein

```



```

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 460 POLYGALACTURONASE.
FT ACT SITE 315 315 PROBABLE. (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED. (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED. (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 460 AA: 49403 MW; 13E99A7A049FAFCE CRC64;

Query Match 45.6%; Score 1091; DB 1; Length 460;
Best Local Similarity 49.3%; Pred. No. 2.4e-69;
Matches 231; Conservative 76; Mismatches 126; Indels 36; Gaps 12;

QY 1 MWIGNSILLIIIFASSIST--CRSNVID--NLFKQYVNDILRQFAHDFQAYLSYLSK 57
D 1 MALKTQILMSFVVVAVFSTSCSGSFGQVNAHLSYVDHVDKESGYNRAVPSY-TD 59
QY 58 NIE-----SNNNIDKVDKNGI-----KVINLVSPFAGKGGKTYDNI 95
D 60 TIEGLKWEELIRPTOLFSSRKLMTI-TGGIATSSAPAKTISVDDFGKKGAD-DTQAF 117
QY 96 EQANNEACSSRTPOFVYVFKKNKYLKQITFSGPCRSISVYKIFGSLASXISDYKDR 155
D 118 VKAKKACSSSGANVLVVPQ-KNVLVRIEFGSPCKSQLTIQITGTEASEDRSTYKDID 176
QY 156 LMIAPDSVQNLVVGSGGTINGNGVWPPSSCKINKSLPCRD-APFALTFWNCKNLKYNL 214
D 177 HMLIFDVQNLVVGPGTINNGNIMWKNCKIKPFCGTAPDPAVTFNCKNLVVKNL 236
QY 215 KSKKAOQIHIFKEBCTVVAANLMMASAKSPNTDGVVNSTOYIOTSDITIGGDCIS 274
D 237 NIQAQOIHVFQNCINQACSLTVLAPEDSPNDGHHVNTQNTITSSVIGTGDCIS 296
QY 275 IVSSGQVQATNITCGFGHGISIGSLGSEAVSVNTVEAKTIGANGVRIKTKWQG 334
D 297 IVSSGQVQATNITCGFGHGISIGSLGSEAVSVNTVEAKTIGANGVRIKTKWGG 356
QY 335 SGQASNTKFLNENQVYKPIIIDQNTCD-RVEPCIQQFSAYQVKNVYENIKTSATKV 393
D 357 SGSAKTNIVQVQVNDVNTPIIIDQNTCDHKTCKCKQKQKSAVQKAVLYQNIKTSASGD 416
QY 394 AIKFDGCTNPFCEGIMENINLVGSEKPSATCKNHNNAEHVTPHC 442
D 417 AITLNGSQSVPCGIVLQSVOL--QNGR--AECNNVQPAYKGVVSPRC 460

RESULT 5
MPA2_CRYUA STANDARD; PRT; 514 AA.
AC PA3212;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
DE (Major pollen allergen Cry j 2) (Cry j II).
OS Cryptometaria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Cupressaceae; Cryptomeria.
NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=95010777; PubMed=7926035;
RA Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,
RA Usui M., Kurimono M.;
RT "Molecular cloning of the second major allergen, Cry j II, from
RL Japanese cedar pollen.";
RL FEBS Lett. 353:124-128(1994).
RP [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94271186; PubMed=8002972;
RA Komiyaama N., Sone T., Shimizu K., Morikubo K., Kino K.;
RT "cDNA cloning and expression of Cry j II the second major allergen of
RT Japanese cedar pollen.";
```

```

RU Biochem. Biophys. Res. Commun. 201:1021-1028(1994).
RN [3]
RP SEQUENCE OF 55-64.
RX MEDLINE=90342988; PubMed=2382797;
RA Sakaguchi M., Inoue S., Tanai M., Ando S., Usui M., Matubasi T.;
RT "Identification of the second major allergen of Japanese cedar
RT pollen.";
RL Allergy 45:309-312(1990).
CC -I- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -I- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
CC -I- ALLERGEN: Causes an allergic reaction in human.
CC -I- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; D37765; BAA07021.1; -.
DR EMBL; D29772; BAA06172.1; -.
DR PIR; JC2498; JC2498.
DR PIR; S48730; S48730.
DR HSSP; P26509; 1BHE.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; PdhL.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00710; PdhL; 5.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
KW Hydrolyase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
KM Amyloplast; Glycoprotein; Allergen.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 45 POTENTIAL.
FT CHAIN 46 433 POSSIBLE POLYGALACTURONASE.
FT PROPEP 434 514 POTENTIAL.
FT ACT SITE 278 278 PROBABLE.
FT CARBOHYD 460 460 N-LINKED. (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED. (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 5 5 F -> L (IN REF. 2).
FT CONFLICT 12 12 V -> L (IN REF. 2).
FT CONFLICT 34 35 DI -> VV (IN REF. 2).
FT CONFLICT 37 37 Q -> K (IN REF. 2).
FT CONFLICT 88 88 K -> N (IN REF. 2).
FT CONFLICT 98 98 N -> S (IN REF. 2).
FT CONFLICT 451 451 K -> E (IN REF. 2).
FT CONFLICT 454 454 G -> R (IN REF. 2).
FT CONFLICT 504 504 M -> I (IN REF. 2).
FT CONFLICT 507 507 R -> C (IN REF. 2).
SQ SEQUENCE 514 AA: 56645 MW; 624611C3FA8D6302 CRC64;

Query Match 32.9%; Score 787.5; DB 1; Length 514;
Best Local Similarity 38.5%; Pred. No. 5.5e-48;
Matches 173; Conservative 76; Mismatches 163; Indels 35; Gaps 9;

QY 8 ILLIIIFASSISTCRSNVIDNLFKQYVNDILRQFAHDFQAYLSYLSKIEBNNNDK 67
D 12 VAKQILIMAAA-----EDQSAQIMLDSIDIO-----YL-----RSNRSLRK 47
QY 68 VD--KNGIKYINLVSPFAGKGGKTYDNIAPFQANNEACSSRTPOFVYVFKKNKYLKQI 124
D 48 VEHRSHALINIFNVEKYGAVGDGHDCTEAFSTMQAAC-KKPSAMLVIPNKKGFVYVNL 106
QY 125 TFGSPCRSSISVYKIFGSLASXISDYKDRRLMIAPDSVQNLVVGSGGTINGNGVWPPS 184
D 107 FENGPCQPHFTFKVDGIIAIVQNPASWKNRINWQFAKLTGFTLMGKGVLDGQKQWAG 166
QY 185 SKC-TINKSLPC--RDAPTALTFWNCKNLKYNLKSRAQOIHIFKEBCTVVAANLMMNA 241
D 167 QCKVWNGREICNDRDRPTAIKFDSTGLIIQGLKLMNSPEFHVIFNGCEGVKIIIGISIT 226
```

```
QY 242 SAKSPNTDGVHVSNTQYIQTSTTIGTDDCSIVSGSQNVQATNTTCGPGHGISIGSIG 301
DB 227 PRDSEPTDIDIDIFASKRPHLQKNTIGTDDCVAGTSSNIVIEDICGPGHGISIGSIG 286
QY 302 SGNSESVSNVNVYENAKITGAENGVRKTMQGGSGQASNKFLNVMQVKKYITIDONY 361
DB 287 RENSREAVSVYHVGAKFIDTQNGRIKTMQGGSGQASHIYENVMISENITLINOQF 346
QY 362 CDREPEICIQFSAVQVKNVYENIKGTSAIKVAIKFDCSTNPFCEGLIMENINLVGSEGK 421
DB 347 CTSASACQORSAVQVQDVTYKIRKTSATMAAITQKCDSDMCKDKIKSDISKTSK 406
QY 422 PSEATCKNVHFN--AEHVTPECTSLSEIS 448
DB 407 --IASCLINDNANGYFSGHVTLPACKNLSPS 433

RESULT 6
PGL1 ARATH STANDARD; PRT; 422 AA.
AC P49062;
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Exopolysaccharuronase clone GBGE184 precursor (EC 3.2.1.67) (ExoPG)
DE (pectinase) (Gallacturan 1,4-alpha-galacturonidase).
GN PG43 OR AT1G02790 OR T14P4.2 OR F22D16.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. C24; TISSUE=flower buds;
RA MEDLINE=99413295; PubMed=10485285;
RX Torik M., Mandaron P., Thomas F., Quigley F., Mache R., Falconet D.;
RT "Differential expression of a polygalacturonase gene family in
RT Arabidopsis thaliana."
RT Mol. Gen. Genet. 261:948-952(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC Torik M., Thomas F., Mache R., Mandaron P., Falconet D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=21016719; PubMed=11130712;
RX Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Brehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Egan P., Feildlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hiltzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetschka I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Millischer U., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson U., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RT Nature 408:816-820(2000).
CC -1- FUNCTION: May function in depolymerizing pectin during pollen
CC development, germination, and tube growth. Acts as an exo-
CC polygalacturonase.
CC -1- CATALYTIC ACTIVITY: (1,4)-alpha-D-galacturonide (N) + H(2)O =
CC (1,4)-alpha-D-galacturonide (N-1) + D-galacturonate.
CC -1- SUBCELLULAR LOCATION: Secreted.
```

```
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; X72291; CA451032.1; -.
DR EMBL; Y16230; CA46127.1; -.
DR EMBL; AC009525; AA02888.1; -.
DR PIR; S34199; S34199.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; Pbh1.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00710; Pbh1; 5.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
DR HydroLase; Glycosidase; Cell wall; signal; glycoprotein;
KM Multigene family.
FT CHAIN 1 31 POTENTIAL.
FT ACT SITE 32 422 EXOPOLYGALACTURONASE CLONE GBGE184.
FT CARBOHYD 264 264 PROBABLE.
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 422 AA; 44430 MW; 17B02E3CE979BA6A CRC64;

Query Match 29.6%; Score 708.5; DR 1; Length 422;
Best Local Similarity 40.8%; Pctd. No. 1,5e-42;
Matches 151; Conservative 66; Mismatches 128; Indels 25; Gaps 11;

QY 75 VINLVSFGAKGDKDXTDNIAFEQAMNACSSRPVQVFNKKNVLYLKQITFGSPCRSSI 134
DB 51 VYDITFGAVGDSSTTFKAFPLNTWVQVCDSPVPAFLVLPKG-TPLAGPTFFAGPCKSKV 109
QY 135 SVKIFGSLSEASSKISDPYKDRRLMIAFDSYQNVLVGGGGITNGQ-VWMPSSCKINKSLP 193
DB 110 TVNVIGTITAT--SGYATPE-WFLFERVDNLVLTGTFHKGGEAWKADGC--GKVVQ 164
QY 194 CRDAPFALTFWCKNKNLKNVNAQOIHKESECTNVVASLUMINAKSKSPRTGVHV 253
DB 165 CNLPPTSLKFRMKNKVEINGISSVNAKAFMFLVKTENNVIQIKLTAPAESPTDQIHL 224
QY 254 SNTQYIQTSTTIGTDDCSIVSGSQNVQATNTTCGPGHGISIGSGNSZAYVSNVT 313
DB 225 SNADNVSIIDSTIATGDDCVSVGRSSNVTVERVICGPGHGLSVGLKRYNEDVSGIH 284
QY 314 VNEAKITGAENGVRKTMQGGSG--GQASNIKFLNVEMQVKKYPIIIDONYCDREPEICIQ 371
DB 285 VNNCTMIENIDNGIRKRTM--GGSDPSKAVDIKFNENIMQSKNPIIIDONYGSGGD---- 339
QY 372 FSAVQVKNVYENIKGTSAIKVAIKFDCSTNPFCEGLIMENINLV--VGSEG---KSE- 424
DB 340 -SQVASDILFKNIRGRTITKDVQVIMKSKSVPCQGVVADVLDVYKTKGKSGSSGG 398
QY 425 ---ATCKNVH 431
DB 399 LVGALCDNAN 408

RESULT 7
PGL1 PRUPE STANDARD; PRT; 393 AA.
AC P48979;
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase).
OS Prunus persica (Peach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
```

```

OC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCB1_TaxID=3760;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Flavocresc;
RX MEDLINE=94302167; PubMed=8029352;
RA "Peach (Prunus persica) endopolylacturonase cDNA isolation and mRNA
RT analysis in melting and nonmelting peach cultivars.";
RL Plant Physiol. 105:225-231(1994)
CC -1- FUNCTION: Acts in concert with the pectinesterase, in the ripening
CC process. Is involved in cell wall metabolism, specifically in
CC polyuronide degradation.
CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76735; CA54150.1; -.
DR PIR; S40123; S40123.
DR HSSP; P26509; 1BHE.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; Pbh1.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00710; Pbh1; 6.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
KW Hydrolyase; Glycosidase; Cell wall; Signal; Fruit ripening;
KW Glycoprotein.
FT SIGNAL 1 25
FT CHAIN 26 393
FT ACT SITE 242 242
FT CARBOHYD 260 260
SQ SEQUENCE 393 AA; 41472 MW; 53DCC6944D22BF9 CEC64;

Query Match
Best Local Similarity 28.8%; Score 687.5; DB 1; Length 393;
Matches 148; Conservative 60; Mismatches 137; Indels 13; Gaps 8;

QY 77 NVLSFGAKGDKTYDNIAFEQANNEACSSRTPVOFVVKNNKYLKQTFSGPCR-SSTS 135
DQ 29 NVASLGAADKDTSTRAFLFSAKACASNNPQVITYVAG-TFELRDVVFSGPKNNAIT 87
QY 136 VKIFGSLAASKISDYK---DRLMIAFDSYQNLVVGGGGTINGNGQVWVWSSCKINSL 192
DQ 88 FRINGTIVAP---SDRYVIGNAWVIFPHHNGVTI-SGGLDDQGTALW--ACKACHGE 141
QY 193 PCRDAPALTLFWNCKNLKVNLSKSKNAQOIHKESSCTNVVNASMLTASAKSPNTGVH 252
DQ 142 SCPEGATTLGFSDSNNIYVSGLASINSQMFHIVINDFQNVOMQGVRSRSGNSPNTGIIH 201
QY 253 VSNVQYIQTIIIGTDDCIISYSGSNVQATNITGPRHGHSISGLSGNSRAYSVN 312
DQ 202 VQWSSGVATILSKATGDDCVISIGPSTSNMIEGACGPGHISISGLGKEQEBAGVQNV 261
QY 313 TVNEAKTIIGENGRIKTV--QGSGSOASNIKELNVEMQDVYKPIIIDQNYCDRVEPCIQ 371
DQ 262 TVKVTVTSGTNGRIKSMGSRPSTGPANILFGHATWVNVNPIVIDQHYEPDKKCGCQ 321
QY 372 FSAVQVKNVVENIKGTSATKVAIKPDCSTNFPCEGIIMENINIVSGSKPSEATCKN 429
DQ 322 VSGVQISDVYEDIHGTISATEVAVFDCSPGHPCREIKLEDVKLTYNK-QAASSCSH 378

```

RESULT 8
FGLR_GOSHI

```

ID FGLR_GOSHI STANDARD; PRT; 407 AA.
AC Q39786;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polylacturonase precursor (BC 3.2.1.15) (PG) (Pectinase).
GN G9.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCB1_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Coker 312; TISSUE=pollen;
RX MEDLINE=95161720; PubMed=7858233;
RA John M.E., Petersen M.W.;
RT "Cotton (Gossypium hirsutum L.) pollen-specific polylacturonase
RT mRNA: tissue and temporal specificity of its promoter in transgenic
RT tobacco.";
RL Plant Mol. Biol. 26:1989-1993(1994).
CC -1- FUNCTION: May function in the depolymerization of the pectin in
CC its walls during pollen tube elongation, or in that of the pistil
CC during pollination.
CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -1- TISSUE SPECIFICITY: Pollen.
CC -1- DEVELOPMENTAL STAGE: Appears 12 days before anthesis and maximum
CC levels are seen in pollen on the day of anthesis.
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U09717; AAA82167.1; -.
DR PIR; S52006; S52006.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; Pbh1.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00710; Pbh1; 5.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
KW Hydrolyase; Glycosidase; Cell wall; Signal; Glycoprotein.
KW Glycoprotein.
FT SIGNAL 1 26
FT CHAIN 27 407
FT ACT SITE 244 244
FT CARBOHYD 182 182
FT CARBOHYD 267 267
FT CARBOHYD 272 272
FT CARBOHYD 302 302
FT CARBOHYD 331 331
SQ SEQUENCE 407 AA; 43921 MW; B81B2BC4C312D195 CEC64;

Query Match
Best Local Similarity 27.9%; Score 667; DB 1; Length 407;
Matches 140; Conservative 69; Mismatches 150; Indels 6; Gaps 5;

QY 67 KVDKKGKIVNVLSFGAKGDKTYDNIAFEQANNEACSSRTPVOFVVKNNKYLKQTF 126
DQ 23 KVQSDARDV--VAFAKAGADKDTLSRPFDAWEACASVTPVIVLPKG-TYLLSKNL 79
QY 127 SGPCRSSISVYKIFSLAASKISDYKCRRLMIAFDSYQNLVVGGGGTINGNGQVWVWSSC 186
DQ 80 EGPKKARIEINVOGTIOAPADPSAFKDPN--WVRYSVENFPMFGGJFDGGSTAYEKNT 138
QY 187 KINKSLPCRDAPALTLFWNCKNLKVNLSKSKNAQOIHKESSCTNVVNASMLTASAKSP 246
DQ 139 CENRPRSK-LPVNIRDFLTNALIDITSKDSKLFHINVFACKNITLERKIEAPDESP 197

```


Db	159	ANGNKAFIDGGQSLAMKANDCA--KTKGKNSLPINRFLGILNSKINSITSTNSKLFHMN	216
QY	226	FESCTNNVASNLMITNIAASAKSPMTDGVHNSVTOYIQISDTITIGDDPCISIVSSQNVQAT	285
Db	217	ILMCKNITLSDIGDIAPELPSLNTDGIHIGRSQVMNIGAKIKITGGDCAVSIIGDTENMLIVE	276
QY	286	NITCGPGHGISTGSGSGNSEAVSVNTVNAEKITGAENGVRKTMQGG-SQASNIKFL	344
Db	277	NVECGPGHGISTGSGRYRNEQPVKGVYRKCLIKMTDGVAKIKITPGSPPGIASNILEE	336
QY	345	NVEMQDVKKYPIIIDQNYCDRVEPCIQQFSAYGVKNVYENIKGTSATKYAIKFDGCTNFP	404
Db	337	DITMDNVSLPVLIDQEYCPYGHCKAGVPSQVLTSLVTIKIGIKGTSATKYAVKLMCKSGVP	396
QY	405	CEGIIMENINLV--GESGKPSAATKCV	430
Db	397	CTNIALSDINLVHNKEG-PAVSASGNI	423

RESULT 10	PGLR_GOSBA	STANDARD;	PRT;	407 AA.
ID	PGLR_GOSBA			
AC	Q39766;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Polygalacturonase precursor (EC 3.2.1.15) (PG (Pectinase) . G9).			
GN	Gossypium barbadense (Sea-island cotton) (Egyptian cotton) .			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.			
OX	NCBI_TaxID=3634;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pollen.			
RX	MEDLINE=95161720; PubMed=7858233;			
RA	John M.E., Petersen M.W.;			
RT	"Cotton (Gossypium hirsutum L.) pollen-specific polygalacturonase			
RT	mRNA: tissue and temporal specificity of its promoter in transgenic			
RL	Plant Mol. Biol. 26:1989-1993 (1994) .			
CC	-I- FUNCTION: May function in the depolymerization of the pectin in			
CC	its walls during pollen tube elongation, or in that of the pistil			
CC	during pollination.			
CC	-I- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-			
CC	galactosiduronic linkages in pectate and other galacturonans.			
CC	-I- TISSUE SPECIFICITY: Pollen (By similarity).			
CC	-I- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch) .			
CC	or send an email to license@isb-sib.ch) .			
CC	-----			
DR	EMBL; U09805; AAA58322.1; -			
DR	InterPro; IPR000743; Glyco_hydro_28.			
DR	InterPro; IPR006626; PBH1.			
DR	Pfam; PF007295; Glyco_hydro_28; 1.			
DR	SMART; SM00710; PBH1; 5.			
DR	PROSITE; PS00502; POLYGALACTONASE; 1.			
KW	Hydrolase; Glycosidase; Cell wall; Signal; Glycoprotein.			
FT	SIGNAL	1	26	POTENTIAL.
FT	CHAIN	27	407	POLYGALACTONASE.
FT	ACT SITE	244	244	BY SIMILARITY.
FT	CARBOHYD	182	182	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD	272	272	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD	302	302	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD	331	331	N-LINKED (GLCNAC . . .) (POTENTIAL).

SQ SEQUENCE 407 AA; 43943 MW; EBE842A77158CCF0 CRC64;

Query Match	27.9%	Score 666	DB 1	Length 407
Best Local Similarity	38.4%	Pred. No. 1.4e-39		
Matches 140		Conservative 69	Mismatches 150	Indels 6
				Gaps 5

[illegible]

```

RESULT 11
PGLT_MAIZE
ID_PGLT_MAIZE STANDARD; PRT; 410 AA.
AC P35339;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Expolysgalacturonase precursor (EC 3.2.1.67) (ExoPG) (Pectinase)
DE (Galacturan 1,4-alpha-galacturonidase) .
DE PG2C.
OS Zea mays (Maize) .
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
CX NCBI_TaxID=4577;
[1]
RM SEQUENCE FROM N.A.
RP STRAIN=cv. Missouri 17; TISSUE=leaf;
RC MEDLINE=3164262; PubMed=8433375;
RA Characterate A., Martin W., Outigley F., Maché R.;
RT "Characterization of a multigene family encoding an
exopolysgalacturonase in maize.";
RL J. Mol. Biol. 229:797-801(1993) .
CC -1- FUNCTION: May function in depolymerizing pectin during pollen
development, germination, and tube growth. Acts as an exo-
polygalacturonase.
CC -1- CATALYTIC ACTIVITY: {(1,4)-alpha-D-galacturonide}(N) + H(2)O =
CC {(1,4)-alpha-D-galacturonide}(N-1) + D-galacturonate.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Pollen.
CC -1- DEVELOPMENTAL STAGE: Late stages of pollen development.
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X66422; CAA47052.1; -.

DR PIR; S30067; S30067.

DR MaizEDB; 65847; -.

DR InterPro; IPR000743; Glyco_hydro_28.

DR InterPro; IPR006626; Pbh1.

DR Pfam; PF00295; Glyco_hydro_28; 1.

DR SMART; SM00710; Pbh1; 5.

DR PROSITE; PS00502; POLYGALACTURONASE; 1.

KW Hydrolyase; Glycosidase; Cell wall; Signal; Glycoprotein;

KW Multigene family.

FT SIGNAL 1 22

FT CHAIN 23 410

FT ACT SITE 256 256

FT CARBOHYD 89 89

FT CARBOHYD 246 246

FT CARBOHYD 349 349

FT CARBOHYD 387 387

FT CARBOHYD 410 410

SEQUENCE 410 AA; 43296 MW; 5BD1310F588AE7CA CRC64;

Query Match 27.2%; Score 650.5; DB 1; Length 410;

Best Local Similarity 36.6%; Pred. No. 1,7e-38;

Matches 131; Conservative 75; Mismatches 143; Indels 9; Gaps 6;

QY 77 NVLSFGAKGDKTYDNIAEFOAWNEACSSRTPVQVFPKKNLYLKQITFSGPCRSSISV 136

DB 42 DITKLASGNGKTDSTKAVQEMASACGGTGKTLIPKG-DLVPLNFTGCKSDVIT 100

QY 137 KTFGSLAASSKISDYDRIRMIAFDSVQNLVVGSGGTINGNG-QVWPPSCCKINKSLPCR 195

DB 101 QVNGNLLATTDLSQYDHGHWIEILRVDMVLVTGKGLDQGPVAVMSKXSCV--KCYDCK 158

QY 196 DAPFALTFNNCKNKLKNNLKSRAQOIHKFESCTVAVASNLMTNASASPNTDGVHVN 255

DB 159 IIPNSLVMDPVNNGVSGITLLNSKPFHNMKCKMGLKDVAVTPAGDSPTNDGIHMD 218

QY 256 TOYIQTSDPTLIGTDDCISIVSGSQNVQATNITCGPHGISISGLASGNSAVVSVTVN 315

DB 219 SSGVITTNVIVIGVDDCISIGPBTSKVNTGVTGCGHGISISGLRYDEKDYTLINVK 278

QY 316 EAKIIAENGVRIRKTNQGSNG--QASNIKFLNEMQDVKYPPIIIDONYCDRVEPCIQOPS 373

DB 279 DCLTKKTANGVRKAYEDASVLTASKIHENIKMDSGYPPIIIDMKYCNKLTANGAS 338

QY 374 ANQVKNVVENIKGTATKATKATPCDSTNPFCEGIIMENINL--VESGKPSAATCKN 429

DB 339 KTVVQVTFKNTIGTSTPEAVMLCTAKIPCTGVMTDDVNIKYSGSTNKK-TMAVCKN 395

RESULT 12

PGRL_OENOR STANDARD; PRT; 362 AA.

AC P24548;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Exopolysialacturonase (EC 3.2.1.67) (ExoPG) (Pectinase) (Galacturan 1,4-alpha-galacturonidase) (Fragment).

OS Oenothera organensis (Evening primrose).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Myrtales; Onagraceae; Oenothera.

OC NCBI_TaxID=3945;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MUNZ (EVERSON); TISSUE=Pollen;

RX MEDLINE=93005658; Pubmed=2152116;

RA Brown S.M., Crouch M.L.;

RT "Characterization of a gene family abundantly expressed in Oenothera organensis pollen that shows sequence similarity to

RT polygalacturonase.";

RL Plant Cell 2:263-274(1990).

CC -!- FUNCTION: May function in depolymerizing pectin during pollen development, germination, and tube growth. Acts as an exo-polygalacturonase.

CC -!- CATALYTIC ACTIVITY: {(1,4)-alpha-D-galacturonide} (N) + H(2)O = {(1,4)-alpha-D-galacturonide} (N-1) + D-galacturonate.

CC -!- TISSUE SPECIFICITY: Pollen tubes growing through the style during pollination.

CC -!- DEVELOPMENTAL STAGE: Accumulates late in pollen development and/or during germination and tube growth.

CC -!- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.

DR PIR; UQ0992; UQ0992.

DR InterPro; IPR000743; Glyco_hydro_28.

DR InterPro; IPR006626; Pbh1.

DR Pfam; PF00295; Glyco_hydro_28; 1.

DR SMART; SM00710; Pbh1; 4.

DR PROSITE; PS00502; POLYGALACTURONASE; 1.

KW Hydrolyase; Glycosidase; Cell wall; Multigene family.

KW NON_TER 1 1

FT ACT SITE 202 202

FT CARBOHYD 140 140

FT CARBOHYD 192 192

FT CARBOHYD 195 195

FT CARBOHYD 225 225

SEQUENCE 362 AA; 38206 MW; 0AD6F45B5518B947 CRC64;

Query Match 26.4%; Score 630; DB 1; Length 362;

Best Local Similarity 37.9%; Pred. No. 4e-37;

Matches 129; Conservative 69; Mismatches 136; Indels 6; Gaps 6;

QY 94 AFEQAWNEACSSRTPVQVFPKKNLYLKQITFSGPCRSSISVTFGSLAASSKISDYKD 153

DB 5 ALTTAKKACASAPETIIVPRG-NEAVGLITLFGPKRSSIGHQGTLLKAPADPSKIKG 63

QY 154 RRLMLTAFDSVQNLVVGSGGTINGNG-QVWPPSCCKINKSLPCGDATALTFFNNCKNLVN 212

DB 64 LG-WINLANKIDLTITFGGVFQDQGSAAVQNDCHNGPT-CTLSMNLRLVAVTNSILR 121

QY 213 NLKSKRAQOIHKFESCTVAVASNLMTNASASPNTDGVHVSNTQYIQTSDPTLIGTDDC 272

DB 122 DVTTLDSKPFHNVIGCKNLTPERFKISAETSINTDGHIGSGDVNIINTIKTGDDC 181

QY 273 ISIVSGSQNVQATNITCGPHGISISGLASGNSAVVSVTVNEAKIIAENGVRIRKTN 331

DB 182 ISLDGDSKNNITNITICGPHGISVGLRKYNEBSVGVYVNGCTITSGONGVRIRKTN 241

QY 332 QGSGQASNIKFLNEMQDVKYPPIIIDONYCDRVEPCIQGFAVQYKAVVYENIKGTSAT 391

DB 242 KSEPGASPMHRQDITMNSVGTPIIIDOGCCPYNOCTAEVSSVSKLSIFKNIKGTSTT 301

QY 392 KVAIKFDCSTNPFCEGIIMENINLV-GEQKPSAATCKNV 430

DB 302 KEAVKLVCKSPFCNGVELADIDLTYSKGKGPATSVCENI 341

RESULT 13

PGRL_MAIZE STANDARD; PRT; 410 AA.

AC P26216;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Exopolysialacturonase precursor (EC 3.2.1.67) (ExoPG) (Pectinase) (Galacturan 1,4-alpha-galacturonidase).

GN PGI AND PG2 AND PG3 AND PG6 AND PG14.

OS Zea mays (Maize).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACCAD clade; Panicoideae; Andropogoneae; Zea.

OC NCBI_TaxID=4577;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-37.

RX STRAIN=cv. Missouri 17; TISSUE=Pollen;

RX MEDLINE=92032781; PubMed=1932692;
 RA Nigret M.F., Dubald M., Mandaron P., Mache R.;
 RT "Characterization of pollen polygalacturonase encoded by several cDNA
 clones in maize.";
 RL Plant Mol. Biol. 17:1155-1164(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Missouri 17; TISSUE=leaf;
 RX MEDLINE=93164262; PubMed=8433375;
 RA Barakate A., Martin W., Outigley F., Mache R.;
 RT "Characterization of a multigene family encoding an
 expopolgalacturonase in maize.";
 RL J. Mol. Biol. 229:797-801(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. B73, and cv. W22;
 RX MEDLINE=93004490; PubMed=1391780;
 RA Allen R.L., Lonsdale D.M.;
 RT "Sequence analysis of three members of the maize polygalacturonase
 RT gene family expressed during pollen development.";
 RL Plant Mol. Biol. 20:343-345(1992).
 RN [4]
 RP SEQUENCE OF 1-306 FROM N.A.
 RX MEDLINE=94035141; PubMed=8106080;
 RA Allen R.L., Lonsdale D.M.;
 RT "Molecular characterization of one of the maize polygalacturonase
 RT gene family members which are expressed during late pollen
 development.";
 RL Plant J. 3:261-271(1993).
 CC -1- FUNCTION: May function in depolymerizing pectin during pollen
 development, germination, and tube growth. Acts as an exo-
 polygalacturonase.
 CC -1- CATALYTIC ACTIVITY: {(1,4)-alpha-D-galacturonide} (N) + H(2)O =
 {(1,4)-alpha-D-galacturonide} (N-1) + D-galacturonate.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Pollen.
 CC -1- DEVELOPMENTAL STAGE: Late stages of pollen development.
 CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X57627; CAA40850.1; -;
 DR EMBL; X57628; CAA40851.1; ALT_INIT.
 DR EMBL; X57743; CAA40910.1; -;
 DR EMBL; X57743; CAA40910.1; -;
 DR EMBL; X57743; CAA40910.1; -;
 DR EMBL; X65844; CAA46679.1; -;
 DR EMBL; X65844; CAA46679.1; -;
 DR EMBL; X65845; CAA46680.1; -;
 DR EMBL; X62384; CAA44248.1; -;
 DR EMBL; X62385; CAA44249.1; -;
 DR EMBL; X66692; CAA47234.1; -;
 DR PIR; S18570; S18570.
 DR PIR; S25824; S25824.
 DR PIR; S30064; S30064.
 DR MaizeDB; 25864; -;
 DR InterPro; IPR000743; Glyco_hydro_28.
 DR InterPro; IPR006626; PH1.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR SMART; SM00710; Pbh1; 5.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 KW Hydrolase; Glycosidase; Cell wall; Signal; Glycoprotein;
 KW Multigene family.
 FT CHAIN 1 22
 FT ACT SITE 23 410 EXOPOLYGALACTURONASE.
 FT CARBOYD 256 256
 FT CARBOYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 359 359 A -> V (IN PG614).
 SQ SEQUENCE 410 AA; 43443 MW; 0A6E779644D818BA CRC64;
 Query Match 26.2%; Score 627; DB 1; Length 410;
 Best Local Similarity 34.4%; Pred. No. 7.6e-37;
 Matches 134; Conservative 77; Mismatches 157; Indels 22; Gaps 8;
 QY 66 DKVDKNGIKV-----INVLSPGAKGDKTYDNIAPFOAMNEAGSSRTYQVAVPEKX 117
 DB 23 EKESKSGIDAKASPGGSFDITKLGASGNGRTDSKVAQVEMASACGCTGKQTIIIPKG- 81
 QY 118 NYLKKQITFGSPGCRSSISVKIFGSLPSSKISDPKRLMTAFDSVQXLYVGGGSTINGN 177
 DB 82 DFLVQGLNFTGPKGDVYIQVDSNLATTDLSQKDHQNMTEILRVNDLVITGKNLDSQ 141
 QY 178 G-QVWPPSSCKINKSLPGRDAPALTFWNCNKLKVNINKSKNAQOIHIFESCTNVVANS 236
 DB 142 GPVWSSKNSC--TKYVQCKILPNSLVMDPVNNGVSGVTLLNSKFPHNMWRCKDMLIKD 199
 QY 237 LMINASAKSPYTDGVHVSNTQYIQISPTIITGDPDCTISVSGSNVQAKTNTTCGPHGIS 296
 DB 200 VVTAPGDSPTNDGILHMGDSGITTNTVIGVGDCTISGTSKVNITGVYCGGHGIS 259
 QY 297 IGSIGSGNSGSEAYVSNVTFNNEAKITGAENGVRIRIKTWOGSG--QASNIKFLVNMODYKYP 354
 DB 260 IGSIGRKYKDEVDYDINVKDCTLAKTMEGVRIKAYEDASVLTYSKIHYENIKHDSANP 319
 QY 355 IIDONTCDRVEPCITQPSAVQVKNVYENIKTSATKVAIKFDCSTNPFEGITMENIN 414
 DB 320 IFIMKXCPNKLCTVANGSKVTVMDVTFKNITGTSSTPEAVSLICTAKVPCGVMTDOVN 379
 QY 415 L--VGESGKPSKACXKVNHFNAEHVPHC 442
 DB 380 VEYSGTNKK-TMAIC-----TNAGSTKGC 403
 RESULT 14
 ID PGIS_MAIZE STANDARD; PRT; 410 AA.
 AC P35338;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Expopolgalacturonase precursor (EC 3.2.1.67) (ExoB) (Pectinase)
 GN (Galacturon 1,4-alpha-galacturonidase).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAP clade; Panicoideae; Andropogoneae; Zea.
 CX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Missouri 17; TISSUE=Leaf;
 RX MEDLINE=93164262; PubMed=8433375;
 RA Barakate A., Martin W., Outigley F., Mache R.;
 RT "Characterization of a multigene family encoding an
 expopolgalacturonase in maize.";
 RL J. Mol. Biol. 229:797-801(1993).
 CC -1- FUNCTION: May function in depolymerizing pectin during pollen
 development, germination, and tube growth. Acts as an exo-
 polygalacturonase.
 CC -1- CATALYTIC ACTIVITY: {(1,4)-alpha-D-galacturonide} (N) + H(2)O =
 {(1,4)-alpha-D-galacturonide} (N-1) + D-galacturonate.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Pollen.
 CC -1- DEVELOPMENTAL STAGE: Late stages of pollen development.
 CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

Search completed: July 21, 2004, 17:06:51
Job time : 23.6574 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 21, 2004, 17:02:32 ; Search time 69.4797 Seconds

(without alignments)
2075.310 Million cell updates/sec

Title: US-10-018-604-2

Perfect score: 2390
Sequence: 1 MWIGNSILLIIPASSIS.....VTPHCTSLSEDEALYNY 457

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SEPREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282	53.6	463	10	Q9M6S2 actinidia c
2	1204.5	50.4	444	10	Q94B15 vitis vinif
3	1185.5	49.6	428	10	Q944B5 glycine max
4	1185	49.6	438	10	Q95FB7 arabidopsis
5	1149	48.1	433	10	Q84Y29 arabidopsis
6	1130	47.3	434	10	Q95FDY7 brassica na
7	1128	47.2	461	10	Q81246 cucumis mel
8	1127	47.2	431	10	Q23147 arabidopsis
9	1115.5	46.7	426	10	Q22935 arabidopsis
10	1113	46.6	433	10	Q42399 brassica na
11	1108	46.4	458	10	Q43063 prunus pers
12	1106	46.3	433	10	Q8VXT3 brassica ra
13	1103	46.2	433	10	Q42636 brassica na
14	1073	44.9	460	10	Q8GTP8 pyrus commu
15	1061.5	44.4	335	10	Q65886 rhus idaeu
16	1034.5	43.3	405	10	Q84ZP3 oryza sativ

17	1022.5	42.8	422	10	Q84ZP5 oryza sativ
18	1014	42.4	444	10	Q84ZN9 oryza sativ
19	988.5	41.4	444	10	Q84M40 oryza sativ
20	925	38.7	452	10	Q9M7D3 lycopersico
21	914	38.2	444	10	Q94AJ5 arabidopsis
22	898.5	37.6	459	10	Q9SSC2 arabidopsis
23	840.5	35.2	468	10	Q9CAL5 arabidopsis
24	825.5	34.5	435	10	Q9SLP3 cucumis sat
25	789	33.0	514	10	Q8H987 cryptomeria
26	788.5	33.0	514	10	Q8H989 cryptomeria
27	788	33.0	457	10	Q9ZUR7 arabidopsis
28	787.5	32.9	434	10	Q9FXC1 arabidopsis
29	787.5	32.9	514	10	Q8H988 cryptomeria
30	764	32.0	507	10	Q9FY19 juniperus a
31	749.5	31.4	1161	10	Q9L0D1 arabidopsis
32	741.5	31.0	491	10	Q9FWX5 arabidopsis
33	736	30.8	423	10	Q81245 cucumis mel
34	734	30.7	503	10	Q94I21 oryza sativ
35	730	30.5	392	10	Q22311 lycopersico
36	728	30.5	405	10	Q22818 arabidopsis
37	728	30.5	443	10	Q9SM63 glycine max
38	727	30.4	392	10	Q40135 lycopersico
39	726	30.4	392	10	Q96487 lycopersico
40	719.5	30.1	398	10	Q8GTP7 pyrus commu
41	718.5	30.1	395	10	Q22610 lycopersico
42	718	30.0	351	10	Q84N55 prunus pers
43	718	30.0	435	10	Q9LYJ5 arabidopsis
44	715.5	29.9	397	10	Q9ZT65 lycopersico
45	715	29.9	392	10	Q80559 arabidopsis

ALIGNMENTS

RESULT 1

ID	Q9M6S2	PRELIMINARY;	PRT;	463 AA.
AC	Q9M6S2;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Polygalacturonase A.			
GN	PGA.			
OS	Actinidia chinensis (Kiwi) (Yangtiao).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			
OC	Ericales; Actinidiaceae; Actinidia.			
OX	NCBI_Taxid=3625;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=pericarp;			
RX	MEDLINE=20252524; PubMed=10794531;			
RA	Wang Z.Y., MacRae F.A., Wright M.A., Bolitho K.M., Rose G.S.,			
RA	Atkinson R.G.;			
RT	"Polygalacturonase gene expression in kiwifruit: relationship to fruit			
RT	softening and ethylene production.";			
RL	Plant Mol. Biol. 42:317-328(2000).			
CC	-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES			
CC	(POLYGALACTURONASES).			
DR	EMBL; AF152758; AAF71160.1; -			
DR	GO; GO:0005618; C:cell wall; IEA.			
DR	GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.			
DR	GO; GO:0004650; F:polygalacturonase activity; IEA.			
DR	GO; GO:0005975; P:carbohydrate metabolism; IEA.			
DR	InterPro; IPR000743; Glyco_hydro_28.			
DR	InterPro; IPR006626; PDB.			
DR	InterPro; IPR000408; Reg_ch_conds.			
DR	Pfam; PF00295; Glyco_hydro_28; 1.			
DR	SMART; SM00710; PDB1; 4.			
DR	PROSITE; PS00502; POLYGALACTURONASE; 1.			
DR	PROSITE; PS00626; RCCL_2; 1.			
KW	Cell wall; Glycosidase; Hydrolase.			
SQ	SEQUENCE 463 AA; 50352 MW; 7EED7EEFE54988FA CRC64;			

Query Match 53.6%; Score 1282; DB 10; Length 463;
 Best Local Similarity 54.2%; Pred. No. 6.9e-76;
 Matches 256; Conservative 72; Mismatches 114; Indels 30; Gaps 8;

QY 1 MTIGRN--SILLIIIFASSISITCRSNVIDNLFKQVYNDLLEOFBFAHFGAYLSTLSKN 58
 1 MALGRHFQFQFVITLLPFIITLPGTSAVHD-----PRHDYHLE-BYGVDFKAYSYITTI 55
 QY 59 IESNNNDKVDKNGI-----KVIIVLSFGAKDGKTYNDIAEQWANE 101
 56 GDNFPGSSMSHENDIFGLRKVDYGMDRVLDASKTYNVDDFGAKDGDSD-DYKAFKAKMKA 114
 QY 102 ACSSTRPVQVFPKKNKYLLKQITFGSPCRSSISVKIFGSLFASAKISDY-KDRRLMTAF 160
 DB 115 VCSSTSSAVLLVPQ-KYLVLPITFGSPCKSDLTQIYGTLEASDDSDYSKQGRHMLVF 173
 QY 161 DSVQNLVVGGGTNGNGQVWMPSSCKINKSLPCRDAPALTLPWNCNKLKYNLKSRYAQ 220
 DB 174 DSVQNLVVGGGTNGNGQVWMPSSCKINKSLPCRDAPALTLPWNCNKLKYNLKSRYAQ 220
 QY 221 QIHFKFSCITNVASNLMINASAKSPNTDGVHVSNTQYIQTSDTITIGTGDCTISVGSQ 280
 DB 234 QIHVSFPCNVAVQASNLMTVPSPNSPTDGIHVTGQINIHSSCVIEFGDCISIVSGGQ 293
 QY 281 NVQATNITCGPHGISISLGSSEAYSVNVTVEAKIIGANGVRIKTMQSGSGQASN 340
 DB 294 KVRVNDITCGPHGISISLGSSEAYSVNVTVEAKIIGANGVRIKTMQSGSGQASN 340
 QY 341 IKFLVEMQDVKYPYIIIDQNYCDREVPICIQFSAVQYKNVYENIKGTSATKVAIKFDCS 400
 DB 354 IKFQNVEMHNVENPIIIDQNYCDQDKPCQEGSSAVQYKNVYENIKGTSATKVAIKFDCS 413
 QY 401 TNPCEGIIEMNTLVGSGKPSKSEATCKNHFNNAEHVTHCTLSLESDBA 452
 DB 414 KRPECQGIIVEDVDLEIGGAARALCNVVELSETGVVSPHCPRE---GEEBA 462

RESULT 2
 ID 094B15 PRELIMINARY; PRT; 444 AA.
 AC 094B15;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Polylacturonase Pgl.
 OS Vitis vinifera (Grape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Vitaceae; Vitis.
 CX NCBI_Taxid=29760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Shiraz.
 RA Numan K.J., Davies C., Robinson S.P., Fincher G.B.;
 RT "Characterization of cell wall modifying enzyme activities and their
 RT corresponding cDNAs during grape berry development.",
 RL Planta 0:0-0(2001)
 CC -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 CC EMBL; AY043233; AA81876.1; -;
 DR GO; GO:0005618; C:cell wall; IEA.
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0004650; F:polylacturonase activity; IEA.
 DR GO; GO:0005975; F:carbohydrate metabolism; IEA.
 DR InterPro; IPR000743; Glyco_hydro_28.
 DR InterPro; IPR006626; Pbh1.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR SMART; SM00710; Pbh1, 4.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 DR PROSITE; PS00626; RCG1_2; 1.
 KW Cell wall; Glycosidase; Hydrolase.

SQ SEQUENCE 444 AA; 47419 MW; D94259F4792F132 CMC64;
 Query Match 50.4%; Score 1204.5; DB 10; Length 444;
 Best Local Similarity 61.5%; Pred. No. 7.7e-71;
 Matches 228; Conservative 53; Mismatches 87; Indels 3; Gaps 3;

QY 73 IKVINLVSPFGAKGKTYNDIAFEQWANEACSSRPVQVFPKKNKYLLKQITFGSPCRS 132
 DB 75 VKMNVNVTYGAAGKSGSDATF-AFKAKAKACSSRQSV-LVFPKKNKYLLKQITFGSPCRS 132
 QY 133 SISVKIFGSLFASAKISDY-KDRRLMTAFDSVQNLVVGGGTNGNGQVWMPSSCKINKS 191
 DB 133 SITVQIYGTVOASIDRAVSNNDMTWMLIFENVQNLAVQGGTNGNGKTMWNSCKVNYD 192
 QY 192 LPCRDAPALTLPWNCNKLKYNLKSRYAQQIHFKFSCITNVASNLMINASAKSPNTDGV 251
 DB 193 LPCRDAPALTLPWNCNKLKYNLKSRYAQQIHFKFSCITNVASNLMINASAKSPNTDGI 252
 QY 252 HVSNTQYIQTSDTITIGTGDCTISVSGSNVQATNITCGPHGISISLGSSEAYSVN 311
 DB 253 HVSNTQYIQTSDTITIGTGDCTISVSGSNVQATNITCGPHGISISLGSSEAYSVN 312
 QY 312 VTVNEAKIIGANGVRIKTMQSGSGQASNIKFLVEMQDVKYPYIIIDQNYCDREVPICIQ 371
 DB 313 VTVNGATLSGTNGVRIKTMQSGSGQASNIFQNVEMHNVENPIIIDQNYCDQDKPCSKS 372
 QY 372 FSAVQYKNVYENIKGTSATKVAIKFDCSTNPCEGIIEMNTLVGSGKPSKSEATCKVH 431
 DB 373 SSAVQYKNVYENIKGTSATKVAIKFDCSTNPCEGIIEMNTLVGSGKPSKSEATCKVH 432
 QY 432 FNNAEVTHPHC 442
 DB 433 VTMWGDVSPNC 443

RESULT 3
 ID 094B15 PRELIMINARY; PRT; 428 AA.
 AC 094B15;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Denilence-related endopolylacturonase.
 GN SDPG.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eufosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 CX NCBI_Taxid=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Williams 82.
 RA Christiansen L.C., Dai Degang F., Ulvskov P., Borkhardt B.;
 RT "Examination of the denilence zone in soybean pods and isolation of a
 RT denilence-related endopolylacturonase gene.",
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 (POLYGALACTURONASES).
 CC EMBL; AF34714; ALU30418.1; -;
 DR GO; GO:0005618; C:cell wall; IEA.
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO; GO:0004650; F:polylacturonase activity; IEA.
 DR GO; GO:0005975; F:carbohydrate metabolism; IEA.
 DR InterPro; IPR000743; Glyco_hydro_28.
 DR InterPro; IPR006626; Pbh1.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR SMART; SM00710; Pbh1, 6.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 KW Cell wall; Glycosidase; Hydrolase.
 SQ SEQUENCE 428 AA; 45830 MW; F553DBAC29762CD CRC64;

Query Match 49.6%; Score 1185.5; DB 10; Length 428;
 Best Local Similarity 51.5%; Pred. No. 1.3e-69;

```

RESULT 4
Q9SFB7      PRELIMINARY;          PRT;          438 AA.
AC      Q9SFB7;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-JUN-2003 (TREMBLrel. 23, Last annotation update)
DE      Putative polygalacturonase.
GN      F17A17.31.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Euxaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
OC      eucotsids II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
[1]
RN      SEQUENCE FROM N.A.
RC      STAYIN=cv. Columbia;
RA      Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA      Roming C.M., Koo H., Fujii C.Y., Utehack T.R., Barnstead M.E.,
RA      Bowman C.L., White O., Nerman W.C., Fraser C.M.;
RA      "Arabidopsis thaliana chromosome III BAC F17A1 genomic sequence.";
RL      Submitted (JAN-2001) to the EMBL/Genbank/DBSP databases
CC      -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
      (POLYGALACTURONASES).
CC      EMBL; AC013483; AAF21207.1; -.
DR      GO: GO:0005618; C:cell wall; IEA.
DR      GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR      GO: GO:0004650; F:polygalacturonase activity; IEA.
DR      GO: GO:0005975; F:carbohydrate metabolism; IEA.
DR      InterPro; IPR000745; Glyco_hydro_28.
DR      InterPro; IPR006626; pblh.
DR      InterPro; IPR00408; Reg_chr_condens.
DR      Pfam; PF00295; Glyco_hydro_28; 1.
DR      SMART; SM00710; Pblh; 4.
DR      PROSITE; PS00507; POLYGALACTURONASE; 1.
DR      PROSITE; PS00626; RCCL_2; 1.
KW      Cell wall; Glycosidase; HydroLase.
SQ      SEQUENCE 438 AA; 48444 MW; ACA6A960403EB37 CRC64;

```

```

RESULT 5
Q8RY29
ID Q8RY29 PRELIMINARY; PRT; 433 AA.
AC Q8RY29;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE A2G41850/T1A17.5.
OS Arabidopsis thaliana (Mouse-ear cress) .
OC Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes, Magnoliophyta; eudicotyledons, core eudicots, rosids;
OC eustroids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
[1]
RN
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.U., Meyers M.C., Banh J.,
RA Bowers L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai Y., Lam E., Lee J.M., Lin J., Miranda M., Natusaka M.,
RA Nguyen M., Onodera C.S., Palm C.U., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriguni M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RT Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC (POLYGLACTURONASES).
CC EMBL; AY078936; AAL84942.1; -.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016788; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; Pbh1.
DR InterPro; IPR00408; Reg_chir_condens.

```



```

Db 168 INGNKTMWQNSCKIDSKPCTKAPALTLYNLNKLNVKRLVRKNAQOIQISIEKCNKE 227
Qy 234 ASNMIVASAKSPYTDGVHSNTQYQISDITIGTDDCISIVSGSNVQAINTTCGPH 293
Db 228 VSNVEITAFGDSPTDGIHITNTQINRVSNDISIEDGTQISIEDGTQLOIFDLTCGPH 287
Qy 294 GISIGISGSGNSEAVSVNTVNEAKIIGAENGVARIKTMQGGSGQASNIKELNVEMODVKY 353
Db 288 GISIGISGDNRSKAYVSGINVDGAKFSESDNGVRIKTYGGSGTAKNIIKFCNIMENAVKN 347
Qy 354 PIIIDONVCRVCEICIOQFSAVQVKNVYENIKTSATKVAIKDCSTNPFCEGIMENT 413
Db 348 PIIIDQYCK-DKCEDQSAVQVKNVYKNIISGTSATDVAITLNCSEKYPCCGIVLENV 406
Qy 414 NLVSGSGPSEATCKNVHPNNAEHVPHCT 443
Db 407 KING-----GTASCKNANVKNQGVSPKCS 431

RESULT 9
ID 022935 PRELIMINARY; PRT; 426 AA.
AC 022935;
DT 01-JUN-1998 (Tremblrel. 05, Created)
DT 01-JUN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Putative polygalacturonase.
GN AT2G41850.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; Pubmed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Beell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Xoo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanhaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Frazer D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGALACTURONASES).
CC EMBL; AC002339; AAC02763.1; -.
DR PIR; H84846; H84846.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; PbH1.
DR InterPro; IPR004081; Reg. chr. condens.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00710; PbH1_5.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
DR PROSITE; PS00626; RCC1_2; 1.
DR Cell wall; Glycosidase; Hydrolase.
SQ SEQUENCE 426 AA; 45924 MW; F756754C0390A80B CRC64;

Query Match 46.7%; Score 1115.5; DB 10; Length 426;
Best Local Similarity 49.8%; Pred. No. 4; de-65;
Matches 222; Conservative 74; Mismatches 117; Indels 33; Gaps 7;

```

```

Qy 6 NSIIILIIIFASSISTGRSVNIDNLFKQVY-----DNILEQEFAPHFOAIVLSKNIIE 60
Db 6 NLVVFELMALMMSWCKASRISPNVYDHSYKRRKSLYR--EDITGRSFRVRSR 63
Qy 61 SNNNIDRVKNGKIVNIVLSEFAGKDGKTYDNIAFEQAMNEACSSRTVPQFVPRKNRYL 120
Db 64 TPT-----TVASVPFGAKDGKTDQTAFFVAMKKAQSSNAGVNLVPRKNTYL 112
Qy 121 LKQTFSGPGRSSIVYIFGSLFSSKISDYKDRRLTAPDSQNLVVGGS--GTINGG 178
Db 113 LKSTQLTGPQNSILITVQIFGLTASQKRSYKDISKIMPDGVNLSVDDGDTGVVGNG 172
Qy 179 QVWPPSSCKINKSLPDCADPALTFFMNCNKLKVNLSKNAQOIHIFESCTNVASNL 238
Db 173 ETWQNGSKRKA-----KALTFYNSKSLIVNKLVRNNAQOIQISIEKSNVQENVV 225
Qy 239 TMSAKSPNTDGVHSNTQYQISDITIGTDDCISIVSGSNVQAINTTCGPHGISIG 298
Db 226 VTAPADSPNTDGIHITNTQINRVSESIIGTDDCISIESGSONVQINDITCGPHGISIG 285
Qy 299 SLGSGNSEAVSVNTVNEAKIIGAENGVARIKTMQGGSGQASNIKELNVEMODVKYPIID 358
Db 286 SLGDNNSKAFVSGVTVGAKLGSNTGVRIKTYGSGGTASNIIFONTQMDVKNPPIID 345
Qy 359 QNYCDRYEPCIQFSAVQVKNVYENIKTSATKVAIKFDCSTNPFCEGIMENTNLVGE 418
Db 346 QDYCDK-SKCTEKSANVQVKNVYRDISGTSASNAITFNKSNKYPCCGIVLDRVNIKG- 403
Qy 419 SKRPSSEATCKNVHPNNAEHVPHCTS 444
Db 404 -----GKATCTNANVVDKGAVALPQCS 425

```

```

RESULT 10
ID 042399 PRELIMINARY; PRT; 433 AA.
AC 042399;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Endo-POLYGALACTURONIDASE.
GN PGAL.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_Taxid=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. TOPAS;
RA Petersen M., Sander L., Child R., van Onckelen H., Ulskov P.,
RA Borkhardt B.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
(POLYGALACTURONASES).
CC EMBL; X98373; CA67020.1; -.
DR EMBL; X98800; CA65072.1; -.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; PbH1.
DR InterPro; IPR004081; Reg. chr. condens.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00710; PbH1_6.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
DR PROSITE; PS00626; RCC1_2; 1.

```



```
RT "Isolation, characterization and expression of one
RL endopolygalacturonase from Brassica rapa."
CC Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases
CC -! SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
   (POLYGALACTURONASES).
DR EMBL; AJ428543; CAD21651.2; -.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR000408; Reg_chir_condens.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
DR PROSITE; PS00626; RCC1_2; 1.
KW Cell wall; Glycosidase; Hydrolase.
SQ SEQUENCE 433 AA; 46638 MW; BDD796980D1C653A CRC64;

Query Match 46.3%; Score 1106; DB 10; Length 433;
Best Local Similarity 50.0%; Pred. No. 2,1e-64;
Matches 222; Conservative 69; Mismatches 125; Indels 28; Gaps 7;

QY 7 SILLIIIFASISSTCSNVID-----DNLFKQYDNLLEQFAHDPQAYLSKNIES 61
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 9 AIFLCVILMLACCOLSSNVDDGYGHEDGSFQT--DSLILKINDDDVLTLSKSDRPTTES 66
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 NNNIDKVDKNGIKVINYLSFGAKGDKGTYPDNIAPFOAMNEACSRTPVQFVPRKKNYLL 121
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 67 S-----TVSVSNFGAKGDKGTDPDTPQAFKKAMKACSTNGVTFLLIPGKTYLL 114
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 KQITFSGPCRSSISVKIFGSLKSSKISDY-KDRRLMIADSVQNLVVGSG--GTINGNG 178
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 115 KSIRFRGPCKSLSPFQILGTLASATKRSIDYGNDRKMHLLIEDVNNLSIDGSAQIVGNG 174
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 179 QVWMPSSCKINKSLPCPDATLALTFMCKLKNLNSKNAQOQHIFESCTNVASNLML 238
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 175 KIMWQNSCKIDKSKPCPKATLALTLYNLKNLNVKNLEVRNAQOQIISIEKCNVDDVKNVX 234
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 239 INASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQANTTCGPHGISIG 298
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 235 ITAPGDSFNTDGHIVATKIRISNSDITGDDCISIEDSQNVQINDLTGCPHGHSIG 294
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 299 SLGSGNSEAVSVNTVNEAKIIGAENGVRIKITWOGSGQASNIKFLVEMQDVKYPPIID 358
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 295 SLGDGNSKAVVSGINVDGATLSETDNGVRIRIKTYOGSGGTAKNIKFQIRNDVKNPPIID 354
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 359 QNYCDRVEPCIQOFSAVQVNNVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVE 418
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 355 QNYCDK-DKCEQOESAVQVNNVYRNIGTSATDVAIMFNCVVKYPCQGIIVLEVNINIKG- 412
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 419 SGKPSKATCKNVHFNNAEHTPHC 442
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 413 ----GRASCKNVNVKDKGTGSPKC 432
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 13
Q42636 PRELIMINARY; PRT; 433 AA.
ID 042636
AC 042636;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, last annotation update)
DE Polygalacturonase precursor (EC 3.2.1.15).
OS Brassaica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid2; Brassicales; Brassicaceae; Brassica.
OC NCBI_Taxid=3708;
OX RN
RA SEQUENCE FROM N.A.
RA Jenkins E.S., Paul W., Coupe S.A., Bell S., Davies L., Roberts J.A.;
RT "Characterization and expression of an mRNA encoding polygalacturonase
```

```
RT during pod development in oilseed rape (Brassica napus L.).";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases
CC -! SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
   (POLYGALACTURONASES).
DR EMBL; Z49971; CAA90272.1; -.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; PBH1.
DR InterPro; IPR00408; Reg_chir_condens.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00710; PBH1; 6.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
DR PROSITE; PS00626; RCC1_2; 1.
KW Cell wall; Glycosidase; Hydrolase; Signal.
FT SIGNAL
FT CHAIN 1
SQ SEQUENCE 433 AA; 46670 MW; 39BE624426D26739 CRC64;

Query Match 46.2%; Score 1103; DB 10; Length 433;
Best Local Similarity 49.8%; Pred. No. 3.2e-64;
Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

QY 7 SILLIIIFASISSTCSNVID-----DNLFKQYDNLLEQFAHDPQAYLSKNIES 61
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 9 AIFLCVILMLACCOLSSNVDDGYGHEDGSFQT--DSLILKINDDDVLTLSKSDRPTTES 66
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 NNNIDKVDKNGIKVINYLSFGAKGDKGTYPDNIAPFOAMNEACSRTPVQFVPRKKNYLL 121
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 67 S-----TVSVSNFGAKGDKGTDPDTPQAFKKAMKACSTNGVTFLLIPGKTYLL 114
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 KQITFSGPCRSSISVKIFGSLKSSKISDY-KDRRLMIADSVQNLVVGSG--GTINGNG 178
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 115 KSIRFRGPCKSLSPFQILGTLASATKRSIDYGNDRKMHLLIEDVNNLSIDGSAQIVGNG 174
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 179 QVWMPSSCKINKSLPCPDATLALTFMCKLKNLNSKNAQOQHIFESCTNVASNLML 238
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 175 KIMWQNSCKIDKSKPCPKATLALTLYNLKNLNVKNLEVRNAQOQIISIEKCNVDDVKNVX 234
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 239 INASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQANTTCGPHGISIG 298
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 235 ITAPGDSFNTDGHIVATKIRISNSDITGDDCISIEDSQNVQINDLTGCPHGHSIG 294
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 299 SLGSGNSEAVSVNTVNEAKIIGAENGVRIKITWOGSGQASNIKFLVEMQDVKYPPIID 358
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 295 SLGDGNSKAVVSGINVDGATLSETDNGVRIRIKTYOGSGGTAKNIKFQIRNDVKNPPIID 354
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 359 QNYCDRVEPCIQOFSAVQVNNVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVE 418
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 355 QNYCDK-DKCEQOESAVQVNNVYRNIGTSATDVAIMFNCVVKYPCQGIIVLEVNINIKG- 412
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 419 SGKPSKATCKNVHFNNAEHTPHC 442
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 413 ----GRASCKNVNVKDKGTGSPKC 432
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14
Q8GTP8 PRELIMINARY; PRT; 460 AA.
ID 08GTP8
AC 08GTP8;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
DE Polygalacturonase precursor (EC 3.2.1.15).
OS Pyrus communis (Pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid2; Rosales; Rosaceae; Maloideae; Pyrus.
OC NCBI_Taxid=23211;
OX RN
RA [1]
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:00:07 ; Search time 2.07709 Seconds
(without alignments)
1360.306 Million cell updates/sec

Title: US-10-018-604-4
Perfect score: 50
Sequence: 1 NGNANGQXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003ds:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	10	2	AAR45389
2	48	96.0	10	2	AAW10696
3	48	96.0	10	3	AAW10696
4	48	96.0	10	4	AAW10696
5	40	80.0	979	5	ABP35664
6	39	78.0	321	4	ABP70839
7	39	78.0	449	2	AAW18210
8	38	76.0	427	4	ABW1028
9	37	74.0	149	3	AAW44604
10	37	74.0	443	4	ABW1160
11	37	74.0	582	6	ADA07883
12	37	74.0	582	6	ABW59482
13	37	74.0	1038	6	ABU15903
14	36	72.0	141	5	ABW52643
15	36	72.0	175	6	ABW52643
16	36	72.0	220	4	ABW71175
17	36	72.0	296	4	ABW55333
18	36	72.0	420	5	ABP73816
19	36	72.0	887	6	ABU21462
20	36	72.0	1018	4	AAU37245
21	36	72.0	1018	4	AAU34301
22	36	72.0	1018	6	ABU18922
23	36	72.0	1284	1	AAW1187
24	36	72.0	2248	4	ABW64494
25	35	70.0	10	5	ABW76722

26	35	70.0	19	2	AAW60371	AAW60371	Vasopress
27	35	70.0	86	5	ABP34678	ABP34678	Human ORF
28	35	70.0	121	2	AAW17484	AAW17484	Human Nix
29	35	70.0	125	6	ABJ25282	ABJ25282	Mouse BAC
30	35	70.0	202	4	ABW1097	ABW1097	Human BAC
31	35	70.0	206	5	ABW33329	ABW33329	Human TS
32	35	70.0	208	6	ABW70357	ABW70357	Human adl
33	35	70.0	218	2	AAW17483	AAW17483	Human Nix
34	35	70.0	219	2	AAW93864	AAW93864	Human TSA
35	35	70.0	219	5	ABW33328	ABW33328	Human TSA
36	35	70.0	219	5	ABW5128	ABW5128	Hypoxia-r
37	35	70.0	232	2	AAW63114	AAW63114	A human a
38	35	70.0	246	4	ABW64176	ABW64176	Drosophila
39	35	70.0	328	5	ABP73597	ABP73597	Candida a
40	35	70.0	411	6	ADB07280	ADB07280	Alloiocec
41	35	70.0	424	6	ADB07282	ADB07282	Alloiocec
42	35	70.0	455	4	ABW1952	ABW1952	Pea bligh
43	35	70.0	463	4	ABW71163	ABW71163	Drosophila
44	35	70.0	495	5	ABG93381	ABG93381	C. albica
45	35	70.0	525	6	ABW6839	ABW6839	Protein e

ALIGNMENTS

RESULT 1	AAW45389	standard; protein; 10 AA.
ID	AAW45389	
AC	AAW45389	
XX		
DT	25-MAR-2003 (revised)	
DT	06-JUL-1994 (first entry)	
XX		
DE	PG beta subunit internal fragment.	
XX		
KW	Beta subunit; PG; polylacturonase; tomato; pathogen.	
OS	Lycopersicon esculentum.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 9 /note="unknown"	
XX		
PN	BP577252-A1.	
PD	05-JAN-1994.	
XX		
PF	07-MAY-1993; 93EP-00303533.	
XX		
PR	08-MAY-1992; 92US-00880915.	
XX		
PA	(UYAR-) UNIV ARIZONA STATE.	
XX		
PI	Delapenna D;	
DR	WPI; 1994-009528/02.	
PT	Transgenic tomato plants contg. lowered levels of polylacturonase	
PT	isoform 1 - for prodn. of tomatoes less susceptible to pathogens and	
XX	requirng less energy to inactivate poly-galacturonase activity.	
XX		
PS	Example; Page 6; 15pp; English.	
CC	N-terminal sequence analysis of purified PG beta-subunit was performed.	
CC	Two internal fragments (AAW45389-90) were used for the construction of	
CC	primers (AAQ54493-95). Two beta-subunit cDNA clones and fragments of 11	
CC	others were sequenced. The consensus DNA sequence obtained is given in	
CC	AAQ54492. (Updated on 25-MAR-2003 to correct PN field.)	
XX		
SQ	Sequence 10 AA;	
Query Match	96.0%; Score 48; DB 2; Length 10;	

Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
| | | | |
Db 1 NNGANGQXV 10

RESULT 2

AAW10696
ID AAW10696 standard; peptide; 10 AA.

AAW10696;

AC 25-MAR-2003 (revised)

DT 09-APR-1997 (first entry)

DE Tomato polygalacturonase isoform 1 beta subunit internal peptide.

KW Tomato; high soluble pectin content; recombinant; antisense; low cost;
transgenic; beta subunit.

OS Lycopersicon esculentum.

Key Location/Qualifiers

FT Misc-difference 9 /note= "unspecified amino acid"

XX US5569831-A.

XX 29-OCT-1996.

PF 11-JUL-1994; 94US-00273538.

PR 08-MAY-1992; 92US-00880915.

PA (UYAR-) UNIV ARIZONA.

PI Dellapenna D;

DR WPI; 1996-496969/49.

PT Transgenic tomato plants with low level of polygalacturonase isoform 1 -
have higher content of soluble pectin(s) and require less heat energy to
eliminate polygalacturonase activity of the fruit.

XX Example 2; Col 15-16; 19pp; English.

CC AAW10696 and AAW10697 are internal peptides of the beta subunit of tomato
polygalacturonase isoform 1 (PG1). The two peptides were generated by
digestion with Lys-C and Gly-C endoproteases respectively. The peptides
were used for the design of degenerate PCR primers used for cDNA library
screening. A cDNA molecule encoding PG1 was identified, this sequence was
used to produce constructs, antisense to the PG1 coding sequence, which
are used to transform a tomato plant cell so as to reduce or eliminate
PG1 activity. There are two isoforms of tomato polygalacturonase (PG),
CC PG1 and PG2. PG1 has a much higher heat stability level (usually at least
by 20 deg. C.) than PG2. By reducing or eliminating PG1 expression,
CC leaving the less heat-tolerant isoform PG2. The process for production of
CC tomato pastes, sauces etc. becomes less costly due to a lower temperature
CC being needed for PG breakdown. A higher viscosity is also produced due to
CC increased soluble pectin content. (Updated on 25-MAR-2003 to correct PF
field.)

XX Sequence 10 AA;

QY Query Match 96.0%; Score 48; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
| | | | |
Db 1 NNGANGQXV 10

RESULT 3

AB19329
ID AB19329 standard; peptide; 10 AA.

AB19329;

DT 19-FEB-2001 (first entry)

DE Polygalacturonase beta-subunit peptide produced after Lys-C digestion.

KW Polygalacturonase beta-subunit; promoter; tomato; fruit; gene expression;
transgenic plant.

OS Lycopersicon esculentum.

Key Location/Qualifiers

FT Misc-difference 9 /note= "unknown amino acid"

XX US6127179-A.

PD 03-OCT-2000.

PF 17-APR-1996; 96US-00632806.

PR 17-APR-1996; 96US-00632806.

XX (DELL/) DELLAPENNA D.

PA (WATS/) WATSON C F.

PI Dellapenna D, Watson CF;

XX WPI; 2000-611059/58.

PT Tomato polygalacturonase beta-subunit promoter, useful for controlling
expression of target genes which affect the quality of developing fruit
without affecting the ripening process.

XX Example; Col 9; 15pp; English.

CC The present sequence represents an internal fragment of the tomato
polygalacturonase beta-subunit. The specification describes a
polygalacturonase beta-subunit promoter. The promoter exhibits both high
expression and high specificity, and provides tissue-preferential and
developmentally regulated expression. The beta-subunit promoter is highly
active in developing tomato fruit from as early as 10 days after
pollination upto the onset of ripening, at which time it becomes
inactive. The fruit specific, developmentally regulated promoter, can
control the expression of target genes which affect the quality of
developing fruit without affecting the ripening process. Target genes
include genes responsible for sugar or starch metabolism, source sink
relation, organic acid balance, flavour components, pathogen resistance,
soluble solids, and water/pH relations. All or part of the promoter may
be used to direct the expression of foreign or endogenous genes in the
sense orientation, or in the antisense orientation, for overexpression or
cosuppression

XX Sequence 10 AA;

QY Query Match 96.0%; Score 48; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
| | | | |
Db 1 NNGANGQXV 10

RESULT 4
AAB48353
ID AAB48353 standard; peptide; 10 AA.

```

XX AA848353;
AC
XX
XX
XX 20-APR-2001 (first entry)
DT
XX
XX Polyalacturonase (PG) enzyme internal Lys-C peptide fragment.
DE
XX
XX Pectin; pectin methyltransferase; PME; polygalacturonase; PG; tomato; PTOM6;
XX food product; yogurt; milk; fruit juice; whey drink; de-esterification.
XX
XX Lycopersicon esculentum.
OS
XX
XX Key Location/Qualifiers
XX FT Misc-difference 9
XX FT /note= "unknown"
XX
XX WO200078982-A1.
XX
XX 28-DEC-2000.
XX
XX 15-JUN-2000; 2000WO-IB000869.
XX
XX 17-JUN-1999; 99GB-00014209.
XX
XX (DANI-) DANISCO AS.
XX
XX Christensen TWIE, Kreiberg JD;
XX
XX WPI; 2001-091573/10.
XX
XX
XX Modifying pectin, for foodstuffs preparation, involves transforming host
XX having pectin methyltransferase (PME) and polygalacturonase (PG) activity by
XX silencing PG activity, to increase PME to PG ratio.
XX
XX Example 2, Page 42; 78pp; English.
XX
XX The invention provides a new method for modifying pectin that involves
XX providing a host having pectin methyltransferase (PME) activity and
XX polygalacturonase (PG) activity, transforming the host by silencing PG
XX activity to provide an increased PME to PG ratio, preparing a PME extract
XX from the transformed host, and using the PME extract to modify pectin. A
XX PME modified pectin is useful for foodstuffs preparation, and to impart
XX an increased functionality to food products such as yogurt, milk/fruit
XX juice and whey drinks. PME is useful to reduce the number of ester groups
XX in a pectin in a block-wise manner, and to de-esterify two or more
XX adjacent galacturonic acid residue of a pectin on at least substantially
XX all of the pectin chains. The present sequence represents an internal
XX peptide fragment of the PG enzyme, used for designing primers
XX
XX
XX Sequence 10 AA;
SQ
XX
XX Query Match 96.0%; Score 48; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred.No.0.1;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NNGANGGXV 10
XX |||||
XX 1 NNGANGGXV 10
XX
XX Db
XX
XX RESULT 5
XX ABP35664
XX ID ABP35664 standard; protein; 979 AA.
XX
XX AC ABP35664;
XX
XX 24-JUL-2002 (first entry)
XX
XX
XX Fungal ZBC protein sequence #90.
XX
XX Secondary metabolite; fungus; zinc binuclear cluster protein;
XX antibacterial; beta-lactam; anti-hypercholesterolaemic; lovastatin;
XX mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergotamine;
XX

```

```

XX
XX angiogenesis inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;
XX fungal toxin; cell surface receptor; plant growth regulator; pigment;
XX insecticide; antineoplastic.
XX
XX Unidentified.
XX
XX WO200224865-A2.
XX
XX 28-MAR-2002.
XX
XX 19-SEP-2001; 2001WO-US029288.
XX
XX 19-SEP-2000; 2000US-023564P.
XX
XX (MICR-) MICROBIA INC.
XX
XX Holtzman D, Madden K, Maxon M, Sherman A;
XX
XX WPI; 2002-352005/38.
XX
XX N-PSDB; ABN79853.
XX
XX
XX New method for improving the production of a secondary metabolite e.g.
XX antineoplastic agent, ergot alkaloid from a fungus involves modulation of
XX the expression of at least one zinc binuclear cluster protein gene.
XX
XX Disclosure; SEQ ID NO 192; 49pp + Sequence Listing; English.
XX
XX
XX The invention relates to improving the production of a secondary
XX metabolite by a fungus. This involves modulating the expression of at
XX least one ZBC (zinc binuclear cluster protein) gene in a manner to
XX improve the yield of the secondary metabolite. Methods of the invention
XX may be used for improving the production of the secondary metabolite e.g.
XX antibacterial (such as beta-lactam), an anti-hypercholesterolaemic (such
XX as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A),
XX an ergot alkaloid (such as ergotamine), an angiogenesis inhibitor (such
XX as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds,
XX a fungal toxin, a modulator of cell surface receptor signalling, a plant
XX growth regulator, a pigment, an insecticide, or an antineoplastic
XX compound. The method results in a decrease in fermentor run-time, a
XX decrease in the size of the fermentor required for the production of
XX equivalent amounts of the secondary metabolite, or a decrease in the
XX biomass required for the production, which translates into decreased
XX waste that must be handled in downstream processing. The sequences given
XX in records ABP3575-ABP3722 represent ZBC proteins. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 979 AA;
SQ
XX
XX Query Match 80.0%; Score 40; DB 5; Length 979;
XX Best Local Similarity 100.0%; Pred.No.2.6e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NNGANG 7
XX |||||
XX 190 NNGANG 196
XX
XX Db
XX
XX RESULT 6
XX AAB70839
XX ID AAB70839 standard; protein; 321 AA.
XX
XX AC AAB70839;
XX
XX 06-ANG-2003 (revised)
XX
XX 25-JUN-2001 (first entry)
XX
XX
XX S. halstedii cellulase protein.
XX
XX Cellulase; fusion construct; protein synthesis; papermaking; medicine;
XX glyceroldehyde phosphate dehydrogenase; food industry; feed industry;
XX washing composition; catalyst.
XX

```



```

XX OS Streptomyces halstedii.
XX XX
XX PN DE19946047-A1.
XX XX
XX XX 29-MAR-2001.
XX XX
XX PF 25-SEP-1999; 99DE-01046047.
XX PR 25-SEP-1999; 99DE-01046047.
XX XX (BADI ) BASF AG.
XX PI Althoeffer H, Pompejus M, Revuelta Doval JL, Santos M, Jimenez A,
XX PI Benito R, Santamaria R, Fernandez J,
XX XX
XX DR WPI: 2001-291995/31.
XX DR N-PSDB; AAF61508.
XX PT Modified Ashbya organisms with altered protein production properties,
XX PT useful particularly for production of Streptomyces halstedii cellulase.
XX PS
XX PS Example 1; Fig 1; 12pp; German.
XX CC This invention describes a novel microorganism (A) of the genus Ashbya
XX CC for biotechnical production of proteins (I), containing a genetic
XX CC construct so that the protein-synthesis capacity is altered relative to
XX CC the wild-type species A. gossypii ATCC 10895. The invention also
XX CC describes (1) cellulase gene (Iia), for use in (A), having a sequence,
XX CC from Streptomyces halstedii; (2) promoter (P) and transcriptional
XX CC terminator (T) from the sequence encoding the glyceraldehyde phosphate
XX CC dehydrogenase of A. gossypii; (3) genetic structure containing (Iia) plus
XX CC regulatory regions that provide expression in Ashbya; (4) a vector
XX CC containing (Iia), P and/or T, or the structure of (3), plus signals that
XX CC provide replication in host cells or integration into a cellular genome;
XX CC (5) transformed Ashbya, for production of cellulase, containing the
XX CC structure of (3) in replicable form; (6) producing (1) using (A); (7)
XX CC producing (A) by exchanging the promoter and/or increasing the gene copy
XX CC number, and/or by introducing a cellulase-encoding gene; and (8) (I)
XX CC produced by the method of (6). (A) are used for recombinant production of
XX CC endogenous or exogenous proteins, specifically cellulase, useful in
XX CC papermaking, medicine, food/feed industries, washing compositions and/or
XX CC as catalysts. (A) provide higher protein yields than wild-type cells.
XX CC This sequence represents the S. halstedii cellulase protein which is used
XX CC in the construction of the A. gossypii/S. halstedii genetic construct
XX CC to illustrate the method of the invention. (Updated on 06-AUG-2003
XX CC to correct OS field.)
XX SQ
XX SQ Sequence 321 AA;
XX
XX Query Match 78.0%; Score 39; DB 4; Length 321;
XX Best Local Similarity 75.0%; Pred. No. 1.2e+02;
XX Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NGNGANGQ 8
XX DB 256 NGNGSNGE 263
XX
XX
XX RESULT 7
XX ID AAM18210 standard; protein; 449 AA.
XX XX
XX AC AAM18210;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 29-SEP-1997 (first entry)
XX XX
XX DE Cellulomonas fimi endoglucanase.
XX XX
XX KM Endoglucanase; ceaA gene; cellulase; cellulose; glucose.
XX OS Cellulomonas fimi.
XX PF

```

```

XX XX Key Location/Qualifiers
XX FH Peptide 1..31
XX FT /label= Sig_peptide
XX FT Protein 32..449
XX FT /label= Mat_protein
XX FT Region 143..165
XX FT /note= "Pro-Thr-rich conserved sequence of unknown
XX FT function"
XX
XX XX US5643791-A.
XX EN
XX XX
XX XX 01-JUL-1997.
XX PD
XX XX
XX PF 19-JAN-1994; 94US-00185303.
XX XX
XX XX 07-AUG-1986; 86US-00894326.
XX PR 18-DEC-1990; 90US-00630396.
XX PR 19-JUN-1992; 92US-00902460.
XX XX
XX PA (UYBR-) UNIV BRITISH COLUMBIA.
XX XX
XX PI Kilburn DG, Miller RC, Warren RAJ;
XX XX
XX DR WPI: 1997-350249/32.
XX DR N-PSDB; AAT71711.
XX PT Cloned Cellulomonas fimi ceaA gene - encodes endo:glucanase enzyme, for
XX PT conversion of cellulose to glucose.
XX PS
XX PS Disclosure; Fig 5A-F; 20pp; English.
XX XX
XX CC The amino acid sequence (AAM18210) of a Cellulomonas fimi endo-1,4- beta-
XX CC glucanase (EC-3.2.1.4) was deduced from an isolated DNA fragment
XX CC (AAT71711) contg. the ceaA gene. The enzyme is capable of hydrolysing
XX CC beta-1,4-glycosidic linkages in the interior of a cellulose molecule,
XX CC converting the cellulose to glucose. The C-terminal region of the enzyme
XX CC is crucial for activity; deletion of 12 amino acids from that end results
XX CC in loss of all activity. Recombinant endoglucanase can be expressed in
XX CC host cells, esp. E. coli, utilizing vectors incorporating the ceaA gene.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX SQ
XX SQ Sequence 449 AA;
XX
XX Query Match 78.0%; Score 39; DB 2; Length 449;
XX Best Local Similarity 75.0%; Pred. No. 1.7e+02;
XX Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NGNGANGQ 8
XX DB 381 NGNGSNGE 388
XX
XX
XX RESULT 8
XX ID ABB61028 standard; protein; 427 AA.
XX XX
XX AC ABB61028;
XX XX
XX DT 26-MAR-2002 (first entry)
XX DT
XX DE Drosophila melanogaster polypeptide SEQ ID NO 9876.
XX XX
XX KM Drosophila; developmental biology; cell signalling; insecticide;
XX KM pharmaceutical.
XX OS Drosophila melanogaster.
XX XX
XX PN W0200171042-A2.
XX PD
XX PD 27-SEP-2001.
XX XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX PF

```

XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI; 2001-656860/75.
DR N-PSDB; ABL05131.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
PS Disclosure; SEQ ID NO 9876; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AAB37737-
CC AAB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 427 AA:

Query Match 76.0%; Score 38; DB 4; Length 427;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NNGANGQXY 10
:|||||
DB 250 HNGANGNHV 259

RESULT 9
AAY44604
ID AAY44604 standard; protein; 149 AA.
XX
AC AAY44604;
XX
DT 07-APR-2000 (first entry)
XX
DE Maize MLO5 protein.
XX
XX MLO5 protein; ZmMLO5; chromosome 3; mutation; recessive allele;
KM disease resistance; pathogen; anti-infective; antipathogenic; antiviral;
KM antifungal; antihelminthic; anti-arthropod; maize; HvMLO1.
XX
XX
OS Zea mays.
XX
XX
FH Key Location/Qualifiers
FT Region 1..37
FT Domain 38..61
FT /label= Intervening_region
FT /label= Transmembrane_helix
FT Region 62..149
FT /label= C-terminal_region
XX
XX MO200001722-A1.
XX
XX 13-JAN-2000.
XX
XX 07-JUL-1999; 99WO-US015255.
XX
XX 07-JUL-1998; 98US-0091875P.
XX
XX (PION-) PIONEER HT-BRED INT INC.
XX

PI Briggs SP, Simmons CR;
XX
XX MPI; 2000-137296/12.
DR N-PSDB; AAZ49563.
XX
PT Creating or enhancing disease resistance in plants by modulating plant
PT resistance genes expression.
XX
XX Claim 1b; Page 70; 89pp; English.
XX
XX The present sequence is the maize MLO5 (ZmMLO5) protein, which exhibits
CC sequence homology to barley MLO1 (HvMLO1) sequence. MLO5 protein is
CC encoded by a mutation-induced recessive allele MLO5, located on
CC chromosome 3, that confers resistance to plant pathogens. Expression of
CC native MLO genes in plants can be altered by transforming them with a DNA
CC construct comprising the mutated MLO gene. Decreasing the expression or
CC activity of native MLO protein leads to enhanced resistance of plants
CC against pathogens such as fungi, virus, nematodes and insects. MLO gene
CC can be used to isolate related sequences from other plants and as
CC molecular markers used in breeding programs aimed at improving disease
CC resistance. MLO protein has anti-infective, antipathogenic, antiviral,
CC antifungal, antihelminthic and anti-arthropod activity
XX
SQ Sequence 149 AA:

Query Match 74.0%; Score 37; DB 3; Length 149;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
:|||||
DB 96 NNGSNG 102

RESULT 10
ABB61160
ID ABB61160 standard; protein; 443 AA.
XX
AC ABB61160;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 10272.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
XX Drosophila melanogaster.
XX
OS
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI; 2001-656860/75.
DR N-PSDB; ABL05263.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
PS Disclosure; SEQ ID NO 10272; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU01840-ABU16175) and the encoded proteins (ABBS7737-
 CC ABBS72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_poc_sequences
 XX
 SQ Sequence 443 AA;

Query Match 74.0%; Score 37; DB 4; Length 443;
 Best Local Similarity 75.0%; Pred. No. 3.6e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGANGANG 8
 Db 111 NGANGANG 118

RESULT 11

ADA07853
 ID ADA07853 standard; protein; 568 AA.

XX ADA07853;

XX 06-NOV-2003 (first entry)

XX Pinctada maxima nacrein.

XX nacre; nacrein; pearl oyster cultivation; pearl formation.

XX Pinctada maxima.

XX US2003027258-A1.

XX 06-FEB-2003.

XX 27-FEB-2002; 2002US-00086510.

XX 02-AUG-2001; 2001US-0310070P.

XX (CHAN/) CHANG F F.

XX (LIH/) LI H.

XX (HSIE/) HSIEH-LI H M.

XX Chang FF, Li H, Hsieh-Li HM;

XX WPI; 2003-596340/56.

XX N-PSDB; ADA07856.

XX New isolated polynucleotide and polypeptide from the nacre gene of the

XX Pinctada margaritifera species, useful for pearl oyster cultivation.

XX Disclosure; Page 24-25; 29pp; English.

XX The invention relates to an isolated polynucleotide that encodes nacrein.

XX The methods and compositions of the present invention are useful for

XX pearl oyster cultivation and pearl formation. The present sequence

XX represents the amino acid sequence of Pinctada maxima nacrein.

QY 1 NGANGANG 7
 Db 387 NGANGANG 393

RESULT 12

ABBS9482
 ID ABBS9482 standard; protein; 582 AA.

XX ABBS9482;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 5238.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABU03585.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

XX Disclosure; SEQ ID NO 5238; 21pp + Sequence listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA

XX sequences (ABU01840-ABU16175) and the encoded proteins (ABBS7737-

XX ABBS72072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_poc_sequences

XX Sequence 582 AA;

QY 1 NGANGANG 7
 Db 440 NGANGANG 446

Query Match 74.0%; Score 37; DB 4; Length 582;
 Best Local Similarity 85.7%; Pred. No. 4.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 13

ABU15903
 ID ABU15903 standard; protein; 1038 AA.

XX ABU15903;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #1430.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Staphylococcus aureus.

XX WO200277183-A2.

XX 03-OCT-2002.
 PD 21-MAR-2002; 2002WO-US009107.
 XX
 PF
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELITRA) PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR N-PSDB; ACA19773.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 43827; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1038 AA;
 SQ
 Query Match 74.0%; Score 37; DB 6; Length 1038;
 Best Local Similarity 60.0%; Pred. No. 8.7e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NNGANGAGQXV 10
 DB 502 DGGKNGQIT 511
 RESULT 14
 ID ABB83327 standard; protein; 141 AA.
 XX
 AC ABB83327;

XX 02-SEP-2002 (first entry)
 DT
 XX
 DE Murine TSAP6 binding protein #1.
 XX
 XX
 XX Neuroprotective; cytoskeletal; TSAP6; apoptosis; cancer; neurodegeneration;
 XX Tumour Suppressor Activated Pathway 6; TSAP6 binding protein; tumour;
 XX murine; human.
 OS
 XX Mus musculus.
 XX WO200252274-A2.
 XX
 XX 04-JUL-2002.
 PD
 XX
 XX 24-DEC-2001; 2001WO-FR004188.
 PF
 XX
 PR 26-DEC-2000; 2000FR-00017027.
 PR 18-SEP-2001; 2001WO-FR002896.
 XX
 XX (MOLE-) MOLECULAR ENGINES LAB.
 XX
 XX Amson R, Telerman A, Passer B;
 PI WPI: 2002-508914/54.
 DR
 XX
 PT Identifying compounds that inhibit binding of TSAP6 to proteins for
 PT treating cancer and neurodegeneration.
 XX
 XX Claim 1; Page 39; 79pp; French.
 XX
 CC The present invention relates to a method for identifying a compound (I)
 CC that inhibits binding of TSAP (tumour Suppressor Activated Pathway) 6 to
 CC a TSAP6 binding protein. The present sequence is one such TSAP6 binding
 CC protein. (I) are useful for modulating tumour reversal and/or apoptosis
 CC for treating cancer or neurodegeneration
 CC
 XX
 XX Sequence 141 AA;
 SQ
 Query Match 72.0%; Score 36; DB 5; Length 141;
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NNGANGAGQXV 10
 DB 27 NNGKNGXGCHV 36
 RESULT 15
 ID ABR52643 standard; protein; 175 AA.
 ABR52643
 XX
 AC ABR52643;
 XX
 DT 20-JUN-2003 (first entry)
 XX
 DE Protein sequence #SEQ ID 151.
 XX
 XX Multi-protein complex; eukaryote; drug target; diagnosis.
 XX
 XX Saccharomycetes cerevisiae.
 OS
 XX EP1258494-A1.
 PN
 XX
 PD 20-NOV-2002.
 XX
 XX 20-DEC-2001; 2001EP-00130253.
 PF
 XX
 PR 15-MAY-2001; 2001EP-00111774.
 PR
 XX
 PA (CELL-) CELLZOME AG.
 XX
 PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;

PI Marzioch M, Schultz JD, Supertl-Furga GD;
XX
DR WPI: 2003-250078/25.
DR N-PSDB; ACC60685.

XX New isolated protein complexes useful for diagnosing a disease or
PT disorder, or as a target for an active agent of a pharmaceutical,
PT preferably a drug target in the treatment or prevention of disease or
PT disorder.

XX
PS Disclosure; SEQ ID NO 151; 17pp + Sequence Listing; English.

CC The invention relates to multiprotein complexes from eukaryotes. Proteins
CC of the invention and DNA sequences encoding them are given in records
CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
CC obtainable by using a protein as a bait and isolating the set of proteins
CC which is attached thereto from cells. Such protein complexes may comprise
CC up to 30 distinct proteins. Protein complexes of the invention are useful
CC for diagnosing a disease or disorder, or as a target for an active agent
CC of a pharmaceutical, preferably a drug target in the treatment or
CC prevention of a disease or disorder. Note: The sequence data for this
CC patent is not represented in the printed specification, but is based on
CC sequence information supplied by the European Patent Office. The complete
CC document is available on CD-ROM

XX
SQ Sequence 175 AA;

Query Match 72.0%; Score 36; DB 6; Length 175;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NGNGANG 7
|||
|||
Db 62 NGNGANG 68

Search completed: July 21, 2004, 17:06:23
Job time : 5.07709 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:04:52 ; Search time 0.620985 Seconds
(without alignments)
831.356 Million cell updates/sec

Title: US-10-018-604-4
Perfect score: 50
Sequence: 1 NNGANGQXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/Backfillset1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	10	1	US-08-273-538A-1
2	48	96.0	10	3	US-08-632-806A-1
3	39	78.0	286	3	US-09-286-691-27
4	39	78.0	286	3	US-09-687-147-27
5	37	74.0	149	4	US-09-347-650-8
6	35	70.0	19	2	US-08-811-492-127
7	35	70.0	232	2	US-08-773-910-1
8	35	70.0	232	4	US-09-199-892-1
9	35	70.0	435	4	US-09-240-639-10
10	35	70.0	574	4	US-09-142-623-13
11	35	70.0	613	4	US-09-508-264A-3
12	35	70.0	1104	4	US-09-268-347-28
13	35	70.0	1104	4	US-09-268-347-34
14	34	68.0	11	6	5178861-8
15	34	68.0	50	1	US-08-178-477B-16
16	34	68.0	157	4	US-09-025-769B-372
17	34	68.0	157	4	US-09-025-769B-373
18	34	68.0	159	3	US-09-027-449-53
19	34	68.0	159	3	US-08-804-444A-53
20	34	68.0	159	3	US-09-026-985-53
21	34	68.0	159	4	US-09-121-952A-53
22	34	68.0	159	4	US-09-234-340A-53
23	34	68.0	163	4	US-09-025-769B-282
24	34	68.0	211	1	US-08-276-852-34
25	34	68.0	211	1	US-08-133-011-16
26	34	68.0	211	1	US-08-322-730A-16
27	34	68.0	211	1	US-08-387-874-16

28	34	68.0	211	1	US-08-899-575-34	Sequence 34, Appl
29	34	68.0	211	1	US-08-899-575-34	Sequence 34, Appl
30	34	68.0	211	2	US-08-383-619-16	Sequence 16, Appl
31	34	68.0	211	3	US-08-907-739-16	Sequence 16, Appl
32	34	68.0	211	4	US-09-729-597-16	Sequence 16, Appl
33	34	68.0	211	5	PCT-US93-08364-16	Sequence 16, Appl
34	34	68.0	211	5	PCT-US95-08743-34	Sequence 16, Appl
35	34	68.0	232	2	US-08-829-876-103	Sequence 103, App
36	34	68.0	232	4	US-09-234-873A-103	Sequence 103, App
37	34	68.0	232	4	US-09-234-873A-103	Sequence 103, App
38	34	68.0	233	2	US-08-829-876-101	Sequence 101, App
39	34	68.0	233	2	US-08-829-876-105	Sequence 105, App
40	34	68.0	233	3	US-08-438-745-15	Sequence 15, Appl
41	34	68.0	233	3	US-08-438-745-15	Sequence 17, Appl
42	34	68.0	233	3	US-09-219-019-15	Sequence 17, Appl
43	34	68.0	233	3	US-09-219-019-17	Sequence 17, Appl
44	34	68.0	233	4	US-09-234-874A-101	Sequence 101, App
45	34	68.0	233	4	US-09-234-874A-105	Sequence 105, App

ALIGNMENTS

RESULT 1
US-08-273-538A-1
; Sequence 1, Application US/08273538A
; Patent No. 5569831
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; TITLE OF INVENTION: Transgenic Tomato Plants with
; TITLE OF INVENTION: Altered Polygalacturonase Isoforms
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nicholas J. Seay
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273, 538A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608/251-5000
; TELEFAX: 608/251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum
; US-08-273-538A-1

Query Match 96.0%; Score 48; DB 1; Length 10;
Best local similarity 100.0%; Pred. No. 0.041;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
|||||||

Db 1 NNGANGQXV 10

RESULT 2

US-08-632-806A-1

Sequence 1, Application US/08632806A

Patent No. 6127179

GENERAL INFORMATION:

APPLICANT: Dellapenna, Dean

APPLICANT: Watson, Colin P.

TITLE OF INVENTION: Gene Promoter for Tomato Fruit

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53703

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/632,806A

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 920214,90123

TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-251-5000

TELEFAX: 608-251-9186

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

ORGANISM: Lycopersicon esculentum

US-08-632-806A-1

Query Match 96.0%; Score 48; DB 3; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.041;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10

Db 1 NNGANGQXV 10

RESULT 3

US-09-286-691-27

Sequence 27, Application US/09286691

Patent No. 6190189

GENERAL INFORMATION:

APPLICANT: Li, Xin-liang

APPLICANT: Ljungdahl, Lars G.

APPLICANT: Chen, Huizhong

TITLE OF INVENTION: Cellulases and Coding Sequences

FILE REFERENCE: 42-96

CURRENT APPLICATION NUMBER: US/09/286,691

CURRENT FILING DATE: 1999-04-05

EARLIER APPLICATION NUMBER: US 60/027,883

EARLIER FILING DATE: 1996-10-04

EARLIER APPLICATION NUMBER: PCT US97/18008

EARLIER FILING DATE: 1997-10-03

NUMBER OF SEQ ID NOS: 29

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 27

LENGTH: 286

TYPE: PRT

ORGANISM: Cellulomonas fimi

US-09-286-691-27

Query Match 78.0%; Score 39; DB 3; Length 286;

Best Local Similarity 75.0%; Pred. No. 38;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANGQ 8

Db 218 NNGSNGE 225

RESULT 4

US-09-687-147-27

Sequence 27, Application US/09687147

Patent No. 6268198

GENERAL INFORMATION:

APPLICANT: Li, Xin-liang

APPLICANT: Ljungdahl, Lars G.

APPLICANT: Chen, Huizhong

TITLE OF INVENTION: Cellulases and Coding Sequences

FILE REFERENCE: 42-96A

CURRENT APPLICATION NUMBER: US/09/687,147

CURRENT FILING DATE: 2000-10-12

PRIOR APPLICATION NUMBER: US 60/027,883

PRIOR FILING DATE: 1996-10-04

PRIOR APPLICATION NUMBER: PCT US97/18008

PRIOR FILING DATE: 1997-10-03

PRIOR APPLICATION NUMBER: 09/286,691

PRIOR FILING DATE: 1999-04-05

NUMBER OF SEQ ID NOS: 29

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 27

LENGTH: 286

TYPE: PRT

ORGANISM: Cellulomonas fimi

US-09-687-147-27

Query Match 78.0%; Score 39; DB 3; Length 286;

Best Local Similarity 75.0%; Pred. No. 38;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANGQ 8

Db 218 NNGSNGE 225

RESULT 5

US-09-347-650-8

Sequence 8, Application US/09347650

Patent No. 6576814

GENERAL INFORMATION:

APPLICANT: Briggs, Steven

TITLE OF INVENTION: Manipulation of M10 Genes to Enhance Disease Resistance

FILE REFERENCE: 5718-42035718/158714

CURRENT APPLICATION NUMBER: US/09/347,650

CURRENT FILING DATE: 1999-07-06

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 8

LENGTH: 149

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: M105

US-09-347-650-8

Query Match 74.0%; Score 37; DB 4; Length 149;

Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
|||||
Db 96 NNGANG 102

RESULT 6
US-08-811-492-127

; Sequence 127, Application US/08811492
; Patent No. 5834247

; GENERAL INFORMATION:

; APPLICANT: COMB, DONALD G.

; APPLICANT: PERLER, FRANCINE B.

; APPLICANT: JACK, WILLIAM E.

; APPLICANT: XU, MING-CUN

; APPLICANT: HODGES, ROBERT A.

; APPLICANT: NOREN, CHRISTOPHER J.

; APPLICANT: CHONG, SHARONG S.C.

; APPLICANT: ADAM, ERIC

; APPLICANT: SOUTHWORTH, MAURICE

; TITLE OF INVENTION: MODIFIED PROTEIN; METHODS OF THEIR

; TITLE OF INVENTION: PRODUCTION AND METHODS FOR PURIFICATION OF TARGET

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 155

; CORRESPONDENCE ADDRESS:

; ADDRESSER: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.

; STREET: 32 TOZER ROAD

; CITY: BEVERLY

; STATE: MASSACHUSETTS

; COUNTRY: USA

; ZIP: 01915

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/811,492

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/580,555

; FILING DATE: 29-DEC-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/496,247

; FILING DATE: 28-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/146,885

; FILING DATE: 03-NOV-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/004,139

; FILING DATE: 09-DEC-1992

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams, Gregory D

; REGISTRATION NUMBER: 30901

; REFERENCE/DOCKET NUMBER: NEB-036C4

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 508-927-5054

; TELEFAX: 509-927-1705

; TELEX:

; INFORMATION FOR SEQ ID NO: 127:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 19 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-811-492-127

Query Match 70.0%; Score 35; DB 2; Length 19;

Best Local Similarity 85.7%; Pred. No. 9.8;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANG 7
|||||
Db 4 NNGANG 10

RESULT 7

US-08-773-910-1

; Sequence 1, Application US/08773910

; Patent No. 5847093

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Goli, Surya K.

; APPLICANT: Hallman, Jennifer L.

; TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS REG

; TITLE OF INVENTION: ULATOR

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: IBM Compatible

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/773,910

; FILING DATE: Herewith

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0184 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 232 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: Genbank

; CLONE: 1715374

; US-08-773-910-1

Query Match 70.0%; Score 35; DB 2; Length 232;

Best Local Similarity 85.7%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANG 7
|||||
Db 49 NNGANG 55

RESULT 8

US-09-199-892-1

; Sequence 1, Application US/09199892

; Patent No. 6307021

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS REG
TITLE OF INVENTION: U1ATOR
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/199,892
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0184 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1715374
US-09-199-892-1

Query Match 70.0%; Score 35; DB 4; Length 232;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGGANG 7
Db 49 NGGANG 55

RESULT 9
US-09-240-639-10
Sequence 10, Application US/09240639
Patent No. 6350447
GENERAL INFORMATION:
APPLICANT: Chadwick, Brian Paul
APPLICANT: Frischauf, Anna-Maria
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS
FILE REFERENCE: 9598-066
CURRENT APPLICATION NUMBER: US/09/240,639
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 455
TYPE: PRT
ORGANISM: P. sativum
US-09-240-639-10

Query Match 70.0%; Score 35; DB 4; Length 455;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGGANG 8
Db 322 NGGANG 329

RESULT 10
US-09-142-623-13
Sequence 13, Application US/09142623
Patent No. 6337201
GENERAL INFORMATION:
APPLICANT: Koji YANAI et al.
TITLE OF INVENTION: -FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF
TITLE OF INVENTION: ISOLATING -FRUCTOFURANOSIDASE GENE, SYSTEM FOR PRODUCING
TITLE OF INVENTION: -FRUCTOFURANOSIDASE, AND -FRUCTOFURANOSIDASE VARIANT
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,623
FILING DATE: September 10, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 98-0989*/LC(WMC)/144
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acid residues
TYPE: Amino acid
STRANDEDNESS: No. 6337201 relevant
TOPOLOGY: No. 6337201 relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Microorganism: Scopulariopsis brevicaulis IF04843
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1..574
IDENTIFICATION METHOD: E
US-09-142-623-13

Query Match 70.0%; Score 35; DB 4; Length 574;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGGANG 7
Db 71 NGGANG 77

RESULT 11
US-09-508-264A-3
Sequence 3, Application US/09508264A
Patent No. 6566111

```
; GENERAL INFORMATION:
; APPLICANT: YAMAI, Koji
; APPLICANT: NAKANE, Akitaka
; APPLICANT: KONO, Toshiaki
; TITLE OF INVENTION: Beta-Fructofuranosidase And Its Gene
; FILE REFERENCE: 2000-0267*/LC/00144
; CURRENT APPLICATION NUMBER: US/09/508,264A
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/JP98/04087
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: JP/245154/1997
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Scopulariopsis brevicaulis IF04843
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (1)...(613)
US-09-508-264A-3
```

```
Query Match          70.0%; Score 35; DB 4; Length 613;
Best Local Similarity 85.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NNGANG 7
    |||||
Db 71 NNGANGSG 77
```

```
RESULT 12
US-09-268-347-28
; Sequence 28, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 1104
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-28
```

```
Query Match          70.0%; Score 35; DB 4; Length 1104;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 NNGANG 7
    |||||
Db 315 NNGANGSG 321
```

```
RESULT 13
US-09-268-347-34
; Sequence 34, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 1104
```

```
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-34
```

```
Query Match          70.0%; Score 35; DB 4; Length 1104;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 NNGANG 7
    |||||
Db 315 NNGANGSG 321
```

```
RESULT 14
517861-8
; Patent No. 517861
; APPLICANT: VERGARA, UTISES, RUIZ, ANDRES, FERREIRA, ARTURO;
; NUSSENZWEIG, RUTH S.; NUSSENZWEIG, VICTOR N.
; TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPTOPEs
; OF CIRCUMPOROZOITE PROTEINS
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,241
; FILING DATE: 22-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,634
; FILING DATE: 26-OCT-1987
; APPLICATION NUMBER: 649,903
; FILING DATE: 12-SEP-1984
; SEQ ID NO: 8
; LENGTH: 11
517861-8
```

```
Query Match          68.0%; Score 34; DB 6; Length 11;
Best Local Similarity 85.7%; Pred. No. 8.1;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 GNGANG 8
    |||||
Db 4 GNGANGQ 10
```

```
RESULT 15
US-08-178-477B-16
; Sequence 16, Application US/08178477B
; Patent No. 5756343
; GENERAL INFORMATION:
; APPLICANT: WU, CARL, CLOS, JOACHIM;
; APPLICANT: WESTWOOD, J. TIMOTHY.; RABINDRAN, SRIDHAR
; TITLE OF INVENTION: CELL STRESS
; TITLE OF INVENTION: TRANSCRIPTIONAL FACTORS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/178,477B
; FILING DATE: 07-JAN-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/617,910
; FILING DATE: 26-NOV-1990
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
```

NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4103051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 50
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-178-477B-16

Query Match 68.0%; Score 34; DB 1; Length 50;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGNGANGQXV 10
Db 18 NGNPANGSNI 27

Search completed: July 21, 2004, 17:09:28
Job time: 1.62099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:08:27 ; Search time 1.73448 Seconds

(without alignments)
1802.048 Million cell updates/sec

Title: US-10-018-604-4

Perfect score: 50

Sequence: 1 NNGNGNGQXV 10

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1285356 seqs, 312560742 residues

Total number of hits satisfying chosen parameters: 1285356

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	80.0	451	US-10-369-493-4148	Sequence 4148, App
2	40	80.0	979	US-09-801-368-284	Sequence 284, App
3	40	80.0	979	US-10-149-310-192	Sequence 192, App
4	37	74.0	187	US-10-425-114-53903	Sequence 53903, A
5	37	74.0	260	US-10-425-114-56925	Sequence 56925, A
6	37	74.0	380	US-10-425-114-44021	Sequence 44021, A
7	37	74.0	452	US-10-437-963-201301	Sequence 201301, A
8	37	74.0	482	US-10-424-599-150886	Sequence 150886, A
9	37	74.0	568	US-10-086-510-5	Sequence 5, Appl1
10	37	74.0	1038	US-10-282-122A-43827	Sequence 43827, A
11	36	72.0	296	US-09-949-029-70	Sequence 70, Appl1
12	36	72.0	376	US-10-425-114-44598	Sequence 44598, A
13	36	72.0	420	US-10-032-585-7653	Sequence 7653, App
14	36	72.0	572	US-10-424-599-145867	Sequence 145867, A
15	36	72.0	763	US-10-156-761-12117	Sequence 12117, A

16	36	72.0	796	US-10-369-493-3508	Sequence 3508, App
17	36	72.0	887	US-10-282-122A-49386	Sequence 49386, A
18	36	72.0	1018	US-09-815-242-5797	Sequence 5797, App
19	36	72.0	1018	US-09-815-242-12838	Sequence 12838, A
20	35	70.0	86	US-09-864-408A-7302	Sequence 7302, App
21	35	70.0	189	US-10-437-963-170736	Sequence 170736, App
22	35	70.0	219	US-10-170-385-217	Sequence 217, App
23	35	70.0	232	US-09-957-295-1	Sequence 1, Appl1
24	35	70.0	328	US-10-033-585-7434	Sequence 7434, App
25	35	70.0	528	US-10-282-122A-64763	Sequence 64763, A
26	35	70.0	558	US-10-282-122A-50632	Sequence 50632, A
27	35	70.0	574	US-09-990-385-13	Sequence 13, Appl1
28	35	70.0	659	US-10-369-493-1663	Sequence 1663, App
29	35	70.0	728	US-10-032-585-7547	Sequence 7547, App
30	35	70.0	972	US-10-369-493-6221	Sequence 6221, App
31	35	70.0	972	US-10-369-493-6222	Sequence 6222, App
32	34	68.0	26	US-10-329-781A-1	Sequence 1, Appl1
33	34	68.0	53	US-10-424-599-149155	Sequence 149155, A
34	34	68.0	80	US-10-424-599-250266	Sequence 250266, A
35	34	68.0	98	US-10-437-963-121796	Sequence 121796, A
36	34	68.0	150	US-09-884-767A-217	Sequence 217, App
37	34	68.0	152	US-09-837-306-197	Sequence 197, App
38	34	68.0	152	US-10-045-674-454	Sequence 454, App
39	34	68.0	152	US-10-045-674-584	Sequence 584, App
40	34	68.0	152	US-10-437-963-169487	Sequence 169487, A
41	34	68.0	158	US-09-056-160B-102	Sequence 102, App
42	34	68.0	158	US-10-234-671-131	Sequence 131, App
43	34	68.0	159	US-09-726-258-53	Sequence 53, Appl1
44	34	68.0	187	US-10-399-883-34	Sequence 34, Appl1
45	34	68.0	187	US-10-183-687-416	Sequence 416, App

ALIGNMENTS

RESULT 1
US-10-369-493-4148
; Sequence 4148, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianning
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4148
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(451)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-4148

Query Match 80.0%; Score 40; DB 15; Length 451;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGNGNGQ 8
|||||
Db 277 NNGNGSQ 284

RESULT 2

US-09-801-368-284
; Sequence 284, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 284
; LENGTH: 979
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-284

Query Match 80.0%; Score 40; DB 9; Length 979;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
DB 190 NNGANG 196

RESULT 3
US-10-149-310-192
; Sequence 192, Application US/10149310
; Publication No. US20040077039A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas
; APPLICANT: Madden, Kevin T.
; APPLICANT: Maxon, Mary
; APPLICANT: Sherman, Amir
; TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
; FILE REFERENCE: 14184-019051
; CURRENT APPLICATION NUMBER: US/10/149,310
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: PCT/US01/29288
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/233,564
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 192
; LENGTH: 979
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-149-310-192

Query Match 80.0%; Score 40; DB 16; Length 979;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7

DB 190 NNGANG 196
RESULT 4
US-10-425-114-53903
; Sequence 53903, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53903
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17333B01_FLI.pep
US-10-425-114-53903

Query Match 74.0%; Score 37; DB 12; Length 187;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
DB 134 NNGANG 140

RESULT 5
US-10-425-114-56925
; Sequence 56925, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56925
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17086H05_FLI.pep
US-10-425-114-56925

Query Match 74.0%; Score 37; DB 12; Length 260;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
DB 207 NNGANG 213

RESULT 6

```
US-10-425-114-44021
; Sequence 44021, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5321)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44021
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700829626_Flt.pep
US-10-425-114-44021
```

```
Query Match          74.0%; Score 37; DB 12; Length 380;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NNGANG 7
    |||:|
Db 230 NNGSNG 236
```

```
RESULT 7
US-10-437-963-201301
; Sequence 201301, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5321)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 201301
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_96690C.1.pep
US-10-437-963-201301
```

```
Query Match          74.0%; Score 37; DB 16; Length 452;
Best Local Similarity 70.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 NNGANGGXV 10
    |||:|
Db 246 NNGGDHGV 255
```

```
RESULT 8
US-10-424-599-150886
; Sequence 150886, Application US/10424599
; Publication No. US20040031072A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 150886
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(482)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107274C.1.pep
US-10-424-599-150886
```

```
Query Match          74.0%; Score 37; DB 12; Length 482;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NNGANG 7
    |||:|
Db 332 NNGSNG 338
```

```
RESULT 9
US-10-086-510-5
; Sequence 5, Application US/10086510
; Publication No. US20030027258A1
; GENERAL INFORMATION:
; APPLICANT: Fang-Tseh (Frank) CHANG et al.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PEARL
; TITLE OF INVENTION: OYSTER CULTIVATION
; FILE REFERENCE: 505493000120
; CURRENT APPLICATION NUMBER: US/10/086,510
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/310,070
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Pinctada maxima
US-10-086-510-5
```

```
Query Match          74.0%; Score 37; DB 14; Length 568;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NNGANG 7
    |||:|
Db 387 NNGSNG 393
```

```
RESULT 10
US-10-282-122A-43827
; Sequence 43827, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olsen, Karl
; APPLICANT: Zyskind, Judith
```

APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.0344
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43827
LENGTH: 1038
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-282-122A-43827

Query Match 74.0%; Score 37; DB 12; Length 1038;
Best Local Similarity 60.0%; Pred. No. 8.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGNGANGQXY 10
Db 502 DNGKNGQYI 511

RESULT 11
US-09-949-029-70
Sequence 70, Application US/09949029
Publication No. US20030134278A1
GENERAL INFORMATION:
APPLICANT: Karpen, G.H.
APPLICANT: Dobie, K.W.
APPLICANT: Kennedy, C.D.
APPLICANT: Velasco, V.M.
APPLICANT: McGrath, T.L.
APPLICANT: Weko, J.
APPLICANT: Patterson, R.W.
TITLE OF INVENTION: Identification of chromosome inheritance modifiers in Drosophila
FILE REFERENCE: 1211.015US1
CURRENT APPLICATION NUMBER: US/09/949,029
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US 60/231,178
PRIOR FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 149
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 70
LENGTH: 296
TYPE: PRT
ORGANISM: Drosophila melanogaster

US-09-949-029-70

Query Match 72.0%; Score 36; DB 10; Length 296;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANG 7
Db 233 NGNGTNG 239

RESULT 12
US-10-425-114-44598
Sequence 44598, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 44598
LENGTH: 376
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 701013187_FLI.pep
US-10-425-114-44598

Query Match 72.0%; Score 36; DB 12; Length 376;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANGQ 8
Db 53 NGSQTNGQ 60

RESULT 13
US-10-032-585-7653
Sequence 7653, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7653
LENGTH: 420
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7653

Query Match 72.0%; Score 36; DB 14; Length 420;
Best Local Similarity 75.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANGQ 8
Db 189 NGGANGGE 196

RESULT 14

US-10-424-599-145867
; Sequence 145867, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 145867
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(572)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_102737C.1.pep
US-10-424-599-145867

Query Match Best Local Similarity 72.0%; Score 36; DB 12; Length 572;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANGCQ 8

Db 249 NNGANGCQ 256

RESULT 15

US-10-156-761-12117
; Sequence 12117, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIRA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12117
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12117

Query Match Best Local Similarity 72.0%; Score 36; DB 14; Length 763;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANGC 7

Db 744 NNGANGC 750

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:04:03 ; Search time 0.642398 Seconds
(without alignments)
1497.381 Million cell updates/sec

Title: US-10-018-604-4
Perfect score: 50
Sequence: 1 NNGANGGXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	90.0	630	2 JQ1670	polysialacturonase
2	42	84.0	539	2 D71260	hypothetical prote
3	40	80.0	534	2 S62572	hypothetical prote
4	40	80.0	979	2 A38792	transcription acti
5	39	78.0	137	2 H70811	hypothetical glyci
6	39	78.0	449	2 A24993	cellulase (EC 3.2.
7	37	74.0	568	2 J07210	molluscan shell ma
8	37	74.0	694	2 AG2353	translation elonga
9	37	74.0	1038	2 H90053	hypothetical prote
10	37	74.0	1519	2 S41525	major ring-forming
11	36	72.0	65	2 I36850	A26U protein - var
12	36	72.0	65	2 T28568	hypothetical prote
13	36	72.0	101	2 H71266	A27U protein - var
14	36	72.0	175	2 S32957	hypothetical prote
15	36	72.0	253	2 F89030	hypothetical prote
16	36	72.0	253	2 T37247	protein C18G1.5 [i
17	36	72.0	288	1 GNBPT4	histone H1.4 - Cae
18	36	72.0	491	2 D67037	gene 9 protein - p
19	36	72.0	693	2 T36717	hypothetical prote
20	36	72.0	693	2 G71302	probable serine/th
21	36	72.0	774	2 T49573	probable tpr prote
22	36	72.0	1284	1 KMVZAI	probable ATP-bindi
23	36	72.0	1538	1 H70846	A-type inclusion p
24	36	72.0	2248	1 D42088	adenylate cyclase
25	35	70.0	219	2 T34523	hypothetical prote
26	35	70.0	262	2 T22489	hypothetical prote
27	35	70.0	363	2 C97298	probable membrane
28	35	70.0	411	2 B97367	hypothetical prote
29	35	70.0	411	2 AC2585	conserved hypothet

30	35	70.0	455	2 S48859	nucleoside triphos
31	35	70.0	520	2 T22835	hypothetical prote
32	35	70.0	525	2 D70878	hypothetical glyci
33	35	70.0	570	2 T49181	cyclophilin-like p
34	35	70.0	601	2 C89451	protein T04G9.6 [i
35	35	70.0	830	2 S54547	PAM protein - yea
36	35	70.0	969	2 F71418	hypothetical prote
37	35	70.0	972	2 T22488	hypothetical prote
38	35	70.0	1018	2 T22318	hypothetical prote
39	35	70.0	1043	2 AH1906	hypothetical prote
40	35	70.0	1279	2 T18312	hypothetical prote
41	34	68.0	213	2 G82723	partition protein
42	34	68.0	245	2 F70787	hypothetical glyci
43	34	68.0	341	1 G64058	UDP-N-acetylmutram
44	34	68.0	362	2 G84526	hypothetical prote
45	34	68.0	424	1 Z3BPFD	coat protein A pre

ALIGNMENTS

RESULT 1
JQ1670
polysialacturonase (EC 3.2.1.15) 1 beta chain precursor - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 21-Jul-2000
C:Accession: JQ1670; P00447
R:Zheng, L.; Heupel, R.C.; DeLaPenna, D.
Plant Cell 4, 1147-1156, 1992
A>Title: The beta subunit of tomato fruit polysialacturonase isoenzyme 1: isolation, char
A:Reference number: JQ1670; PMID:93005745; PMID:1392611
A:Accession: JQ1670
A:Molecule type: mRNA
A:Residues: 1-630 <ZHE>
A:Cross-references: GB:M98466; NID:G170473; PIDN:AAA34181.1; PID:G170480
A:Accession: P00447
A:Molecule type: protein
A:Residues: 109-119,160-171,230-236,243-251,291-293 <ZHE1>
A:Experimental source: fruit, cv. Alisa Craig
C:Comment: This enzyme is a cell wall protein that hydrolyzes pectins.
C:Keywords: cell wall, glycoprotein, glycosidase, hydrolase, polysaccharide degradation
F:1-30/Domain: signal sequence #status predicted <Sto>
F:31-108/Domain: amino-terminal propeptide #status predicted <Sto>
F:109-387/Product: polysialacturonase 1 beta chain #status experimental <MAT>
F:398-630/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:124,142,256,334,369,387/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.0%; Score 45; DB 2; Length 630;
Best Local Similarity 80.0%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANGGXV 10
Db 243 NNGANGGEFV 252

RESULT 2
D71260
hypothetical protein TP0969 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: D71260
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
iron, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; PMID:98332770; PMID:9665876
A:Accession: D71260
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-539 <COI>
A:Cross-references: GB:AE001264; GB:AE000520; NID:G3323278; PIDN:AAC65927.1; PID:G332329

A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0969

Query Match 84.0%; Score 42; DB 2; Length 539;
Best Local Similarity 80.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
DB 105 NNGANGXKV 114

RESULT 3
Hypothetical protein SPAC30D11.14 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C;Accession: T38585; S62572
R;Pearson, D.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z21801
A;Accession: T38585
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-534 <PE2>
A;Cross-references: EMBL:Z67961; NID:g1065887; PIDN:CAA91900.1; PID:g1065901; GSPDB:GN00
A;Experimental source: strain 972n; cosmid c30D11
C;Genetics:
A;Gene: SPAC30D11.14
A;Map position: 1L

Query Match 80.0%; Score 40; DB 2; Length 534;
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANGQ 8
DB 276 NNGANGQ 283

RESULT 4
transcription activator PUT3 - Yeast (Saccharomyces cerevisiae)
A;Accession: A39792
N;Alternate names: protein YKL015W
C;Species: Saccharomyces cerevisiae
C;Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 24-Sep-1999
C;Accession: A39792; S34701; S37828; S37832; S16705
R;Marczak, J.E.; Brandriss, M.C.
Mol. Cell. Biol. 11: 2609-2619, 1991
A;Title: Analysis of constitutive and noninducible mutations of the PUT3 transcriptional
A;Reference number: A39792; MUID:91203881; PMID:2017167
A;Accession: A39792
A;Molecule type: DNA
A;Residues: 1-979 <MAR>
A;Cross-references: GB:X55384; NID:g4251; PIDN:CAA39055.1; PID:g4252
R;Wieman, S.; Voss, H.; Schwager, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grochues
submitted to the EMBL Data Library, July 1993
A;Description: Sequencing and analysis of 51.5 kilobases on the left arm of chromosome X
A;Reference number: S34679
A;Accession: S34701
A;Molecule type: DNA
A;Residues: 1-979 <WIE>
A;Cross-references: EMBL:X74152; NID:g450363; PIDN:CAA52267.1; PID:g393257
R;Wieman, S.; Voss, H.; Schwager, C.; Rupp, T.; Grochues, D.; Sensen, C.; Stegemann, J.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37825
A;Accession: S37828
A;Molecule type: DNA
A;Residues: 1-979 <W12>
A;Cross-references: EMBL:Z28015; NID:g486002; PIDN:CAA81850.1; PID:g486003; MIPS:YKL015W
A;Experimental source: strain S288C

R;Rieger, M.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37832
A;Accession: S37832
A;Molecule type: DNA
A;Residues: 1-825 <RIE>
A;Cross-references: EMBL:Z28015; MIPS:YKL015W
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:PUT3
A;Cross-references: SGD:S0001498; MIPS:YKL015W
A;Map position: 1L
C;Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluster hc
C;Keywords: DNA binding; nucleus; transcription regulation; zinc finger
F;29-65/Domain: GAL4 zinc binuclear cluster homology <GAL4>
F;34-60/Region: zinc finger

Query Match 80.0%; Score 40; DB 2; Length 979;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
DB 190 NNGANG 196

RESULT 5
H70811
Hypothetical glycine-rich protein Rv0832 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: H70811
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garrier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393: 537-544, 1998
A;Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:9825987; PMID:9634230
A;Accession: H70811
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-137 <COL>
A;Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CAA17638.1; PID:e125397;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0832

Query Match 78.0%; Score 39; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNGANGQ 8
DB 119 NNGANGQ 125

RESULT 6
A24993
cellulase (EC 3.2.1.4) precursor - Cellulomonas fimi
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Cellulomonas fimi
C;Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 31-Mar-2000
C;Accession: A24993
R;Wong, W.K.R.; Gerhardt, B.; Guo, Z.M.; Kilburn, D.G.; Warren, R.A.J.; Miller Jr., R.C.
Gene 44: 315-324, 1986
A;Title: Characterization and structure of an endoglucanase gene cema of Cellulomonas fimi
A;Reference number: A24993; MUID:87055249; PMID:3023193
A;Accession: A24993
A;Molecule type: DNA
A;Residues: 1-449 <WON>
A;Cross-references: GB:M15823; NID:g144409; PIDN:AAA23084.1; PID:g144410
A;Note: the amino-terminal sequence of the mature protein (residues 32-62) has been deter-

C:Genetics:
A:Gene: cema
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel
A:Pathway: cellulose degradation
C:Superfamily: Cellulomonas fimi cellulase; bacterial cellulose-binding domain homology
C:Keywords: glycosidase, hydrolase, polysaccharide degradation
F:1-31/Domain: signal sequence #status predicted <Sig>
F:32-449/Product: cellulase #status experimental <Mat>
F:34-135/Domain: bacterial cellulose-binding domain homology <BCB>
F:35-134/Distaffide bonds: #status predicted

Query Match 78.0%; Score 39; DB 2; Length 449;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNGANGQ 8
 |||||:
Db 381 NNGANGNGE 388

RESULT 7
JCT210
molluscan shell matrix protein N66 - Pinctada maxima
C:Species: Pinctada maxima
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: JCT210
R:Kono, M.; Hayashi, N.; Samata, T.
Biochem. Biophys. Res. Commun. 269, 213-218, 2000
A:Title: Molecular mechanism of the nacreous layer formation in Pinctada maxima.
A:Reference number: JCT210; MUID:20160475; PMID:10694502
A:Accession: JCT210
A:Molecule type: mRNA
A:Residues: 1-568 <KON>
A:Cross-references: DDBJ:AB032612
C:Comment: This protein is rich in asparagine and glycine residues, it serves as a produ
Ayers. It is also important in calcification.
C:Keywords: matrix protein

Query Match 74.0%; Score 37; DB 2; Length 568;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNGANG 7
 |||||:
Db 387 NNGANGNG 393

RESULT 8
AG2353
translation elongation factor EF-G [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AG2353
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2353
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-694 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076082.1; PID:g17133519; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: fus
C:Superfamily: translation elongation factor G; translation elongation factor Tu homolog

Query Match 74.0%; Score 37; DB 2; Length 694;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNGANG 7
 |||||:
Db 685 NNGANGNG 691

RESULT 9
H90053
hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: H90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: GB:BA000018; PID:g13702453; PIDN:BA043594.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb

Query Match 74.0%; Score 37; DB 2; Length 1038;
Best Local Similarity 60.0%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NNGANGQXV 10
 |||||:
Db 502 DNGKNGQII 511

RESULT 10
S41525
major ring-forming surface protein precursor - Helicobacter mustelae
C:Species: Helicobacter mustelae
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 28-May-1999
C:Accession: S41525; S77974
R:O'Toole, P.W.; Austin, D.W.; Trust, T.J.
Mol. Microbiol. 11, 349-361, 1994
A:Title: Identification and molecular characterization of a major ring-forming surface p
A:Reference number: S41525; MUID:94224153; PMID:8170397
A:Accession: S41525
A:Molecule type: DNA
A:Residues: 1-1519 <OTO>
A:Cross-references: GB:LL5629; NID:g431303; PIDN:AC36865.1; PID:g431304
A:Accession: S77974
A:Molecule type: protein
A:Residues: 'XX', 50-59 <AUS>
C:Genetics:
A:Gene: hsr
F:1-47/Domain: signal sequence #status predicted <Sig>
F:48-1519/Product: major ring-forming surface protein #status experimental <MAT>

Query Match 74.0%; Score 37; DB 2; Length 1519;
Best Local Similarity 70.0%; Pred. No. 13e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NNGANGQXV 10
 |||||:
Db 886 NNGANGNAV 895

RESULT 11
I36850
A26L protein - variola virus (strain India-1967)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Mar-2001
C:Accession: I36850

R:Binov, V.M.

submitted to GenBank, November 1992

A:Reference number: A36859

A:Accession: I36850

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-65 <BLI>

A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49071.1; PID:g457021

Query Match 72.0%; Score 36; DB 2; Length 65;

Best Local Similarity 85.7%; Pred. No. 7.9;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANG 7

Db 21 NGNGANG 27

RESULT 12

T28568

hypothetical protein A27L - variola major virus

C:Species: variola major virus

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000

C:Accession: T28568

R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin

Nature 366, 748-751, 1993

A:Title: Potential virulence determinants in terminal regions of variola smallpox virus

A:Reference number: Z20488; MUID:94088747; PMID:8264798

A:Accession: I28568

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-65 <MAS>

A:Cross-references: EMBL:L22579; NID:g623595; PIDN:AAA60878.1; PID:g433048

A:Experimental source: strain Bangladesh-1975

Query Match 72.0%; Score 36; DB 2; Length 65;

Best Local Similarity 85.7%; Pred. No. 7.9;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANG 7

Db 21 NGNGANG 27

RESULT 13

H72166

A27L protein - variola minor virus (strain Garcia-1966)

C:Species: variola minor virus

C>Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999

C:Accession: H72166

R:Shchelkunov, S.N.; Tormentin, A.V.; Guttorov, V.V.; Saitonov, P.F.; Massung, R.F.; Lopat

submitted to GenBank, March 1998

A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor

A:Reference number: A72150

A:Accession: H72166

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-101 <SHC>

A:Cross-references: GB:Y16780; NID:g5830555; PIDN:CA854730.1; PID:e1542686; PID:g5830691

A:Experimental source: strain Garcia-1966

C:Genetics:

A:Gene: A27L

Query Match 72.0%; Score 36; DB 2; Length 101;

Best Local Similarity 85.7%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANG 7

Db 57 NGNGANG 63

RESULT 14

S32957

hypothetical protein YBR254C - Yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBR1722

C:Species: Saccharomyces cerevisiae

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-Apr-2002

C:Accession: S32957; S46135

R:Daignon, F.; Bileau, N.; Crouzet, M.; Aigle, M.

Yeast 9, 189-199, 1993

A:Title: The complete sequence of a 19,482 bp segment located on the right arm of chromo

A:Reference number: S29348; MUID:93220397; PMID:8465606

A:Accession: S32957

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-175 <DOI>

A:Cross-references: EMBL:X70529; NID:g1907246; PIDN:CAA49918.1; PID:g296553

R:Aigle, M.; Baclet, M.C.; Barthe, C.; Bileau, N.; Crouzet, M.; Daignon, F.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45940

A:Accession: S46135

A:Molecule type: DNA

A:Residues: 1-175 <ATG>

A:Cross-references: EMBL:Z36123; NID:g536674; PIDN:CA855217.1; PID:g536675; MIPS:YBR254C

C:Genetics:

A:Gene: SGD:TR520

A:Map position: 2R

Query Match 72.0%; Score 36; DB 2; Length 175;

Best Local Similarity 85.7%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANG 7

Db 62 NGNGANG 68

RESULT 15

F89030

protein Cl8G1.5 (imported) - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001

C:Accession: F89030

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_eleg

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: F89030

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-253 <STO>

A:Cross-references: GB:chr_V; PIDN:AA17754.1; PID:g3168854; GSPDB:GN00023; CESP:Cl8G1.5

C:Genetics:

A:Gene: Cl8G1.5

A:Map position: 5

C:Superfamily: histone H1

Query Match 72.0%; Score 36; DB 2; Length 253;

Best Local Similarity 75.0%; Pred. No. 31;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGNGANG 8

Db 115 SGNGANG 122

Search completed: July 21, 2004, 17:08:52

Job time : 1.6424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:02:02 ; Search time 0.342612 Seconds
(without alignments)
1519.798 Million cell updates/sec

Title: US-10-018-604-4
Perfect score: 50
Sequence: 1 NNGANGQXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	84.0	539	Y969_TREPA	O83935 treponema p
2	40	80.0	534	YABE_SCHPO	Q09911 schizosacch
3	40	79.0	979	PUT3_YEAST	P25502 saccharomyc
4	39	78.0	321	GUNI_STRHA	P33682 streptomyc
5	39	78.0	449	GUNA_CELFI	P07984 cellulomona
6	36	72.0	175	TRC0_YEAST	P38334 saccharomyc
7	36	72.0	288	VG03_BPT4	P10927 bacterioph
8	36	72.0	673	PKNX_STRCO	Q9X416 streptomyc
9	36	72.0	1284	ATI_COMPX	P16602 cowpox viru
10	36	72.0	2248	CYAI_DROME	P32870 drosophila
11	35	70.0	163	CYPH_UROFA	O00060 uromycos fa
12	35	70.0	219	NIEPL_MOUSE	Q92257 mus muscul
13	35	70.0	219	NIEPL_HUMAN	Q60238 homo sapien
14	35	70.0	455	NTPA_PEA	P52914 pisum sativ
15	35	70.0	478	TBGB_RETET	P54405 reticulomyc
16	35	70.0	830	PAMI_YEAST	P37304 saccharomyc
17	35	70.0	856	ALB7_AZCVI	Q92599 azotobacter
18	35	70.0	919	Y893_HUMAN	Q94567 homo sapien
19	35	70.0	981	UBI5_HUMAN	Q9Y468 homo sapien
20	35	70.0	981	UBI5_MOUSE	Q87511 mus muscul
21	34	68.0	341	MURB_HAEIN	P44605 haemophilus
22	34	68.0	424	COAA_BPPD	P03661 bacterioph
23	34	68.0	424	COAA_BPM13	P03662 bacterioph
24	34	68.0	498	Y118_MYCTU	O50615 mycobacteri
25	34	68.0	543	Y991_MYCTU	O50630 mycobacteri
26	34	68.0	544	Y966_TREPA	O83932 treponema p
27	34	68.0	547	YCGV_ECO57	O8X583 escherichia
28	34	68.0	603	YD25_MYCTU	Q10637 mycobacteri
29	34	68.0	778	YQ34_MYCTU	P17933 mycobacteri
30	34	68.0	801	Y747_MYCTU	O53810 mycobacteri
31	34	68.0	833	HSP_YEAST	P10961 saccharomyc
32	34	68.0	914	WA22_MYCTU	O06794 mycobacteri
33	34	68.0	955	YCGV_ECOLI	P76017 escherichia

34	34	68.0	957	1	Y278_MYCTU	P56877 mycobacteri
35	34	68.0	1039	1	Y304_TREPA	O83326 treponema p
36	34	68.0	1253	1	DSDP_HUMAN	Q9NZW4 homo sapien
37	34	68.0	1672	1	PMDB_CHIMU	Q9PJY2 chlamydia m
38	34	68.0	2843	1	APC_HUMAN	P25054 homo sapien
39	33	66.0	306	1	HMTX_DROPS	P20822 drosophila
40	33	66.0	373	1	YHX6_YEAST	P38866 saccharomyc
41	33	66.0	378	1	CSP_PLAVB	P08677 plasmodium
42	33	66.0	386	1	CSP_PLASI	Q03110 plasmodium
43	33	66.0	391	1	PER_DROIN	P91613 drosophila
44	33	66.0	613	1	DP25_PYRFU	P81412 pyrococcus
45	33	66.0	832	1	ANR3_HUMAN	P57078 homo sapien

ALIGNMENTS

RESULT 1
Y969_TREPA STANDARD, PRT; 539 AA.
AC O83935;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0969 precursor.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.C.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Knalax H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman U., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete.";
RL Science 281:375-388 (1998).
CC -1- SIMILARITY: BELONGS TO THE TP096X FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the European Bioinformatics Institute and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
CC EMBL, AE001264, AAC65927.1; -.
DR PIR; D71260;
DR TIGR; TP0969; -.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 15
FT CHAIN 16 539
FT FT 1539
SQ SEQUENCE 539 AA; 59650 MW; 626A657AB35A37E CRC64;
Query Match 84.0%; Score 42; DB 1; Length 539;
Best Local Similarity 80.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 1 NNGANGQXV 10
OY |||||
DB 105 NNGANGQXV 114
|||
RESULT 2
YABE_SCHPO STANDARD, PRT; 534 AA.
ID YABE_SCHPO

```

AC Q09911;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C30D11.14c in chromosome I.
GN SPAC30D11.14c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
CX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Seguros J., Feat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brookes K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jørgensen K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Mablett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicikert G., Aert R., Robben J., Grynolpre B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesli D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Punelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,
RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hunt S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerretti L., Love T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipkavski G.V., Uesery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL/ Z67961; CAA91900.1; -
DR PIR/ T38585; S62572
DR GeneDB/ Spombe; SPAC30D11.14c; -
DR InterPro; IPR004087; KH_dom.
DR SMART; SM00322; KH; 1.
DR KX Hypothetical protein.
DR KM SEQUENCE 534 AA; 59035 MW; B38F20B8A20447F7 CRC64;
SO

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proline utilization trans-activator.
GN PUT3 OR YK015W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91203881; PubMed=2017167;
RA Marczak J.E., Brandtiss M.C.;
RT "Analysis of constitutive and noninducible mutations of the PUT3
RT transcriptional activator.";
RL Mol. Cell. Biol. 11:2609-2619(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=94205264; PubMed=8154185;
RA Wiemann S., Voss H., Schwager C., Rupp T., Stegemann J.,
RA Zimmermann J., Grothues D., Sensen C., Erfle H., Hewltt N.,
RA Banerji A., Ansoorge W.;
RT "Sequencing and analysis of 51.6 kilobases on the left arm of
RT chromosome XI from Saccharomyces cerevisiae reveals 23 open reading
RT frames including the FAS1 gene.";
RL Yeast 9:1343-1348(1993).
RN [3]
RP SEQUENCE OF 1-825 FROM N.A.
RA Rieger M.; (MAR-1994) to the EMBL/GenBank/DBJ databases.
RL Submitted
RN [4]
RP STRUCTURE BY NMR OF 30-100.
RX MEDLINE=97448678; PubMed=9303003;
RA Walters K.J., Davie K.T., Reece R.J., Prashne M., Wagner G.;
RT "Structure and mobility of the PUT3 dimer.";
RL Nat. Struct. Biol. 4:744-750(1997).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 31-100.
RX MEDLINE=97448679; PubMed=9303004;
RA Swaminathan K., Flynn P., Reece R.J., Marmorstein R.;
RT "Crystal structure of a PUT3-DNA complex reveals a novel mechanism
RT for DNA recognition by a protein containing a Zn2Cys6 binuclear
RT cluster.";
RL Nat. Struct. Biol. 4:751-759(1997).
CC -----
CC -1- FUNCTION: POSITIVE ACTIVATOR OF THE PROLINE UTILIZATION PATHWAY.
CC BINDS TO THE PROMOTERS OF PUT1 AND PUT2 GENES. RECOGNIZES AND
CC BINDS TO THE DNA SEQUENCE 5'-CGG-N(10)-CCG-3'.
CC -1- SUBUNIT: Binds DNA as a homodimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL/ X55384; CAA39055.1; -
DR EMBL/ X74152; CAA52267.1; -
DR EMBL/ Z28015; CAA81850.1; -
DR PIR/ A39792; A39792.
DR PDB/ 1AUY; 17-SEP-97.
DR PDB/ 1ZME; 16-SEP-98.
DR GeneDB/ T01163; -
DR TRANSFAC; T01163; -
DR SGD; S0001498; PUT3.
DR GO; GO:0005634; C:nucleus; IC.
DR GO; GO:0003704; F:specific RNA polymerase II transcription fa. .; IDA.
DR GO; GO:0045944; P:positive regulation of transcription from P. .; IDA.
DR GO; GO:0006562; P:proline catabolism; IMP.
DR InterPro; IPR007219; Fungal Trans.
DR InterPro; IPR001138; Fungi_Txn.

```


DR Pfam; PF04082; Fungal_trans; 1.
 DR Pfam; PF00172; Zn_clus; 1.
 DR PRINTS; PR00054; FUNGALZNCYS.
 DR SMART; SM00066; GAL4; 1.
 DR PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; 1.
 DR PROSITE; PS0048; ZN2_Cy6_FUNGAL_2; 1.
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein;
 Zinc; Metal-binding; Proline metabolism; 3D-structure.
 FT DNA_BIND 34 60
 FT METAL 37 37 ZINC 1.
 FT METAL 44 44 ZINC 1.
 FT METAL 50 50 ZINC 1 AND 2.
 FT METAL 53 53 ZINC 2.
 FT METAL 60 60 ZINC 2.
 FT METAL 35 40
 FT METAL 41 41
 FT METAL 51 55
 FT TURN 56 57
 FT STRAND 68 72
 FT HELIX 73 96
 FT TURN 97 98
 SQ SEQUENCE 979 AA; 111414 MW; 59FA19EFC79EE0C0 CRC64;
 Query Match 80.0%; Score 40; DB 1; Length 979;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NGNGANG 7
 DB 190 NGNGANG 196
 RESULT 4
 ID GUN1_STRHA STANDARD; PRT; 321 AA.
 AC P33682;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Endoglucanase 1 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 1)
 DE (Cellulase I) (CMCase I) (CEL1).
 GN CEL1.
 OS Streptomyces halstedii.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1944;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-33.
 RC STRAIN=JMS / CECT3310;
 RX MEDLINE=93015685; PubMed=1400190;
 RA Fernandez-Abalos J.M., Sanchez P., Coll-Fresno P.M.,
 RA Villanueva J.R., Perez P., Santamaria R.I.;
 RT "Cloning and nucleotide sequence of celA1, and
 RT endo-beta-1,4-glucanase-encoding gene from Streptomyces halstedii
 RT JMS";
 RL J. Bacteriol. 174:6368-6376 (1992).
 CC -1- FUNCTION: IMPLICATED IN THE MECHANISM OF INDUCTION EXERTED BY
 CC CELLOBIOSE.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose, lichenin and cereal beta-D-glucans.
 CC -1- SIMILARITY: Belongs to cellulase family B (Family 6 of glycosyl
 CC hydrolases).
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-9 is the initiator.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z12157; CAA78145.1; -.
 DR HSSP; P26222; ITML.
 DR InterPro; IPR001524; Glyco_hydro_6.
 DR Pfam; PF01341; Glyco_hydro_6; 1.
 DR PRINTS; PR00733; GHYDRASE6.
 DR PRODOM; PD003733; Glyco_hydro_6; 1.
 DR PROSITE; PS00655; GLYCOSYL_HYDROL_F6_1; FALSE_NEG.
 DR PROSITE; PS00655; GLYCOSYL_HYDROL_F6_2; 1.
 KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 321 ENDOGLUCANASE 1.
 FT ACT_SITE 110 110 BY SIMILARITY.
 FT ACT_SITE 149 149 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 295 295 NUCLEOPHILE (BY SIMILARITY).
 FT DISTFID 112 156 BY SIMILARITY.
 SQ SEQUENCE 321 AA; 33694 MW; DC5E39EFC79EE0C0 CRC64;
 Query Match 78.0%; Score 39; DB 1; Length 321;
 Best Local Similarity 75.0%; Pred. No. 5.9;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NGNGANG 8
 DB 256 NGNGANG 263
 RESULT 5
 ID GUN1_STRHA STANDARD; PRT; 449 AA.
 AC P07964;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Endoglucanase A precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
 DE (Cellulase).
 GN CBA.
 OS Cellulomonas fimi.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Cellulomonadaceae; Cellulomonas.
 OX NCBI_TaxID=1708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8705249; PubMed=3023193;
 RA Wong W.K.R., Gerhard B., Guo Z.M., Kilburn D.G., Warren R.A.J.,
 RA Miller R.C. Jr.;
 RT "Characterization and structure of an endoglucanase gene cbaA of
 RT Cellulomonas fimi";
 RL Gene 44:315-324 (1986).
 RN [2]
 RP DOMAINS.
 RX MEDLINE=90036847; PubMed=2681184;
 RA Gilkes N.R., Kilburn D.G., Miller R.C. Jr., Warren R.A.J.;
 RT "Structural and functional analysis of a bacterial cellulase by
 RT proteolysis";
 RL J. Biol. Chem. 264:17802-17808 (1989).
 RN [3]
 RP DISULFIDE BONDS.
 RX MEDLINE=92104156; PubMed=1761039;
 RA Gilkes N.R., Claessens M., Aebersold R., Hennisat B., Meinke A.,
 RA Morrison H.D., Kilburn D.G., Warren R.A.J., Miller R.C. Jr.;
 RT "Structural and functional relationships in two families of beta-1,4-
 RT glycanases";
 RL Eur. J. Biochem. 202:367-377 (1991).
 CC -1- FUNCTION: The biological conversion of cellulose to glucose
 CC generally requires three types of hydrolytic enzymes: (1)
 CC Endoglucanases which cut internal beta-1,4-glucosidic bonds; (2)
 CC Exocellulohydrolases that cut the disaccharide cellulose
 CC from the nonreducing end of the cellulose polymer chain; (3)
 CC Beta-1,4-glucosidases which hydrolyze the cellobiose and other
 CC short cello-oligosaccharides to glucose.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose, lichenin and cereal beta-D-glucans.
 CC -1- PTM: The linker region (also termed "hinge") may be a potential

```

CC      site for proteolysis.
CC      -1- SIMILARITY: Belongs to cellulase family B (family 6 of glycosyl
CC      hydrolases).
CC      -1- SIMILARITY: contains 1 bacterial-type cellulose-binding (CBD)
CC      domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M15823; AAA23084.1; -.
DR      PIR; A24993; A24993.
DR      HSSP; P07986; 1EXG.
DR      InterPro; IPR001919; Bac_cellose-bind.
DR      InterPro; IPR008965; Cellul bind.
DR      InterPro; IPR001524; Glyco_hydro_6.
DR      Pfam; PF00553; CBM_2; 1.
DR      Pfam; PF01341; Glyco_hydro_6; 1.
DR      PRINTS; PR00733; GHYDRLASE6.
DR      PRODOM; PD003733; Glyco_hydro_6; 1.
DR      SMART; SM00637; CBD_II; 1.
DR      PROSITE; PS00561; CBD_BACTERIAL; 1.
DR      PROSITE; PS00655; GLYCOSYL_HYDROL_F6_1; 1.
DR      PROSITE; PS00656; GLYCOSYL_HYDROL_F6_2; 1.
KM      Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT      SIGNAL 1 31
FT      CHAIN 32 449
FT      DOMAIN 32 138 ENDOGLUCANASE A.
FT      DOMAIN 139 168 CELLULOSE-BINDING.
FT      DOMAIN 438 449 LINKER ("HINGE") (PRO-THR BOX).
FT      ACT_SITE 247 247 CATALYTIC.
FT      ACT_SITE 283 283 BY SIMILARITY.
FT      ACT_SITE 423 423 PROTON DONOR (BY SIMILARITY).
FT      DISULFID 35 134 NUCLEOPHILE (BY SIMILARITY).
FT      DISULFID 248 291
FT      DISULFID 390 426
SQ      SEQUENCE 449 AA; 46731 MW; 67FF887814B3348D CRC64;

Query Match 78.0%; Score 39; DB 1; Length 449;
Best Local Similarity 75.0%; Pred. No. 8.4;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 NGNGANG 8
DB      381 NGNGANG 388

RESULT 6
TR20_YEAST
ID TR20_YEAST STANDARD; PRT; 175 AA.
AC P38334;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE Transport protein particle 20 kDa subunit (TRAPP 20 kDa subunit).
GN TRS20 OR YBR254C OR YBR1722.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RX MEDLINE=98232495; PubMed=9564032;
RA Schacter M., Jiang Y., Barrowman J., Scarpa A., Burston J., Zhang L.,
RT "TRAPP, a highly conserved novel complex on the cis-Golgi that
RT mediates vesicle docking and fusion."
RL EMBL J. 17:2494-2503 (1998).
CC      -1- FUNCTION: TRAPP plays a key role in the late stages of endoplasmic
CC      reticulum to Golgi traffic.
CC      -1- SUBUNIT: Part of multisubunit complex composed of BET3, BET5,
CC      TRS20, TRS23, TRS31, TRS33, TRS65, TRS85, TRS120 and TRS130.
CC      -1- SIMILARITY: Belongs to the TRAPP small subunits family. Sedlin
CC      subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X70529; CAA49918.1; -.
DR      EMBL; Z36123; CAA85217.1; -.
DR      PIR; S32957; S32957.
DR      Germonline; 138797; -.
DR      SGD; S0000458; TRS20.
DR      GO; GO:0030008; C:TRAPP, IDA.
DR      GO; GO:0006888; P:ER to Golgi transport; IMP.
DR      InterPro; IPR006722; Sedlin_N.
DR      Pfam; PF04628; Sedlin N; 1.
KW Transport; Endoplasmic reticulum; Golgi stack.
SQ      SEQUENCE 175 AA; 19700 MW; 82285CAEAC273509 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 175;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 NGNGANG 7
DB      62 NGNGANG 68

RESULT 7
VC09_BPT4
ID VC09_BPT4 STANDARD; PRT; 288 AA.
AC P10927;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Baseplate structural protein Gp9 (Baseplate wedge protein 9).
GN 9.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OC NCBI_TaxID=10665;
RX MEDLINE=99263746; PubMed=2726468;
RA Prilipov A.G., Selivanov N.A., Efimov V.P., Marusich E.I.,
RT "Nucleotide sequences of bacteriophage T4 genes 9, 10 and 11."
RL Nucleic Acids Res. 17:3303-3303 (1989).
RX MEDLINE=20015372; PubMed=10545330;

```

```

RP      FUNCTION.
RX MEDLINE=98232495; PubMed=9564032;
RA Schacter M., Jiang Y., Barrowman J., Scarpa A., Burston J., Zhang L.,
RA Scheltz D., Yates J.R. III, Abellovich H., Ferro-Novick S.;
RT "TRAPP, a highly conserved novel complex on the cis-Golgi that
RT mediates vesicle docking and fusion."
RL EMBL J. 17:2494-2503 (1998).
CC      -1- FUNCTION: TRAPP plays a key role in the late stages of endoplasmic
CC      reticulum to Golgi traffic.
CC      -1- SUBUNIT: Part of multisubunit complex composed of BET3, BET5,
CC      TRS20, TRS23, TRS31, TRS33, TRS65, TRS85, TRS120 and TRS130.
CC      -1- SIMILARITY: Belongs to the TRAPP small subunits family. Sedlin
CC      subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X70529; CAA49918.1; -.
DR      EMBL; Z36123; CAA85217.1; -.
DR      PIR; S32957; S32957.
DR      Germonline; 138797; -.
DR      SGD; S0000458; TRS20.
DR      GO; GO:0030008; C:TRAPP, IDA.
DR      GO; GO:0006888; P:ER to Golgi transport; IMP.
DR      InterPro; IPR006722; Sedlin_N.
DR      Pfam; PF04628; Sedlin N; 1.
KW Transport; Endoplasmic reticulum; Golgi stack.
SQ      SEQUENCE 175 AA; 19700 MW; 82285CAEAC273509 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 175;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 NGNGANG 7
DB      62 NGNGANG 68

RESULT 7
VC09_BPT4
ID VC09_BPT4 STANDARD; PRT; 288 AA.
AC P10927;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Baseplate structural protein Gp9 (Baseplate wedge protein 9).
GN 9.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OC NCBI_TaxID=10665;
RX MEDLINE=99263746; PubMed=2726468;
RA Prilipov A.G., Selivanov N.A., Efimov V.P., Marusich E.I.,
RT "Nucleotide sequences of bacteriophage T4 genes 9, 10 and 11."
RL Nucleic Acids Res. 17:3303-3303 (1989).
RX MEDLINE=20015372; PubMed=10545330;

```

RA Kostyuchenko V.A., Navruzbekov G.A., Kurochkina L.P., Strelkov S.V.,
RA Mesyanzhinov V.V., Rossmann M.G.;
RT "The structure of bacteriophage T4 gene product 9: the trigger for
RT tail contraction.";
RL Structure 7:1213-1222(1999).
CC -1- FUNCTION: Structural component of the baseplate. Connects the long
CC tail fibers to the baseplate and triggers the tail contraction
CC after virus attachment to a host cell.
CC -1- SUBUNIT: Homotrimer.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14192; CAA32395.1; -;
DR EMBL; AF158101; AAD42520.1; -;
DR PIR; S0A082; GNPPT4.
DR PDB; LOEX; 05-OCT-99.
DR InterPro; IPR008987; Gp9.
KW Structural protein; 3D-structure.
FT TURN 12 13
FT HELIX 20 39
FT TURN 40 41
FT HELIX 45 49
FT TURN 56 57
FT TURN 63 64
FT STRAND 68 68
FT TURN 72 73
FT HELIX 74 76
FT STRAND 79 79
FT TURN 81 82
FT STRAND 84 86
FT STRAND 94 98
FT TURN 103 104
FT STRAND 106 110
FT TURN 112 113
FT STRAND 122 126
FT STRAND 129 130
FT TURN 131 132
FT STRAND 137 139
FT TURN 142 143
FT STRAND 144 152
FT STRAND 159 166
FT STRAND 176 180
FT TURN 183 184
FT STRAND 186 193
FT TURN 194 196
FT STRAND 199 208
FT TURN 209 210
FT STRAND 214 224
FT TURN 225 228
FT STRAND 229 239
FT STRAND 251 257
FT TURN 258 259
FT STRAND 260 266
FT STRAND 272 282
SQ SEQUENCE 288 AA; 30997 MW; 80289CEBA25B760 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 288;
Best Local Similarity 60.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGNGANGXV 10
|||:|:
Db 52 NGTGADGQIT 61

RESULT 8
PKNX_STRCO

ID PKNX_STRCO STANDARD; PRT; 673 AA.
AC Q9XA16;
DT 10-OCT-2003 (rel. 42, Created)
DT 10-OCT-2003 (rel. 42, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Probable serine/threonine-protein kinase SC03848 (EC 2.7.1.37).
GN SC03848 OR SC69.18.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000353;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Lake L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)." ;
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -1- SIMILARITY: Contains 4 PASTA domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL939118; CAB45215.1; -;
DR PIR; T36717; T36717.
DR InterPro; IPR005543; PASTA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF03793; PASTA; 4.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00740; PASTA; 4.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding; Repeat; Complete proteome.
FT DOMAIN 11 277
FT DOMAIN 379 445
FT DOMAIN 446 511
FT DOMAIN 512 580
FT DOMAIN 581 649
FT NP_BIND 17 25
FT BINDING 40 40
FT ACT_SITE 138 138
FT ACT_SITE 138 138
SQ SEQUENCE 673 AA; 71488 MW; 42A0615E239722DE CRC64;

Query Match 72.0%; Score 36; DB 1; Length 673;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANG 7
|||:
Db 654 NGNGANG 660

RESULT 9

```

ATI_COWPX
ID ATI_COWPX STANDARD; PRT; 1284 AA.
AC P1602;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE A-type inclusion protein (ATI).
GN ATI.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CPRO6;
RX MEDLINE=88089536; PubMed=2826668;
RA Funahashi S., Sato T., Shida H.;
RT "Cloning and characterization of the gene encoding the major protein
of the A-type inclusion body of cowpox virus.";
RL J. Gen. Virol. 69:35-47(1988).
RN [2]
RP SEQUENCE OF 1-109 FROM N.A.
RX MEDLINE=8811568; PubMed=2828037;
RA Patel D.D., Pickup D.J.;
RT "Messenger RNAs of a strongly-expressed late gene of cowpox virus
contain 5'-terminal poly(A) sequences.";
EMBO J. 6:3787-3794(1987).
CC -1- FUNCTION: MATURE VIRUSES ARE OCCULDED INTO THE ATI, AND IT HAS
BEN ASSUMED THAT SUCH BODIES PROTECT THE VIRUS DURING
DISSEMINATION FROM ANIMAL TO ANIMAL.
CC -1- MISCELLANEOUS: A CHARACTERISTIC FEATURE OF ATI IS THE FORMATION OF
LARGE MASSES WITH NO SURROUNDING MEMBRANES IN THE CYTOPLASM OF
INFECTED CELLS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; D00319; BAA00222.1; -.
DR EMBL; X06343; CAA29650.1; -.
DR PIR; JQ0006; MWVZAL.
DR InterPro: IPR007596; Pox A type, inc.
DR Pfam; PF04508; Pox_A_type_inc; 14.
KM Late protein; Repeat.
FT DOMAIN 611 912 10 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 611 637 1.
FT REPEAT 637 665 2.
FT REPEAT 665 689 3.
FT REPEAT 689 720 4.
FT REPEAT 720 751 5.
FT REPEAT 751 780 6.
FT REPEAT 780 811 7.
FT REPEAT 811 842 8.
FT REPEAT 842 871 9.
FT REPEAT 871 912 10.
SQ SEQUENCE 1284 AA; 150329 MW; F7904C9E1DEB0C12 CRC64;
Query Match 72.0%; Score 36; DB 1; Length 1284;
Best local similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 NANGANG 7
DB 1023 NANGANG 1029
RESULT 10
CYAI_DROME STANDARD; PRT; 2248 AA.
ID CYAI_DROME

```

```

AC P32870;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ca(2+)/calmodulin-responsive adenylyl cyclase (EC 4.6.1.1) (ATP
pyrophosphate-lyase) (rutabaga protein).
GN RUT.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE=92154664; PubMed=1739965;
RA Levin L.R., Han P.-U., Hwang P.M., Feinstein P.G., Davis R.L.,
RT "The Drosophila learning and memory gene rutabaga encodes a
Ca2+/calmodulin-responsive adenylyl cyclase.";
RL Cell 68:479-489(1992).
CC -1- FUNCTION: This is a membrane-bound, calmodulin-sensitive adenylyl
cyclase. Inactivation of this cyclase leads to a learning and
memory defect.
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -1- COFACTOR: Binds 2 magnesium ions per subunit (by similarity).
CC -1- ENZYME REGULATION: Activated by calcium/calmodulin and G protein.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Mushroom bodies of the fly brain.
CC -1- DOMAIN: Composed of two homologous domains.
CC -1- SIMILARITY: Belongs to the adenylyl cyclase class-4/guanylyl
cyclase family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; M81887; AAA28844.1; -.
DR PIR; D42088; D42088.
DR HSSP; P19754; IAWK.
DR FlyBase: FBmn003301; rut.
DR GO; GO:0019933; P:CaM-mediated signaling; NAS.
DR GO; GO:0007625; P:grooming behavior; NAS.
DR GO; GO:0007591; P:molting cycle (sensu Insecta); IGI.
DR GO; GO:0008355; P:olfactory learning; NAS.
DR GO; GO:0045473; P:response to ethanol (sensu Insecta); NAS.
DR InterPro: IPR001054; G cyclase.
DR Pfam; PF00211; guanylate_cyc; 2.
DR SMART; SM00044; CYCC; 2.
DR PROSITE; PS00452; GUANYLATE CYCLASES 1; 2.
DR PROSITE; PS50125; GUANYLATE CYCLASES 2; 2.
KM CAMP biosynthesis; lyase; Calmodulin-binding; Metal-binding;
Magnesium; Repeat; Transmembrane; Glycoprotein.
FT DOMAIN 1 41 41
FT TRANSMEM 42 60 POTENTIAL.
FT TRANSMEM 65 84 POTENTIAL.
FT TRANSMEM 101 115 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 152 174 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT DOMAIN 207 705 POTENTIAL.
FT TRANSMEM 706 726 POTENTIAL.
FT TRANSMEM 730 750 POTENTIAL.
FT TRANSMEM 770 791 POTENTIAL.
FT DOMAIN 792 813 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 814 834 POTENTIAL.
FT TRANSMEM 842 867 POTENTIAL.
FT TRANSMEM 868 888 POTENTIAL.
FT DOMAIN 889 2248 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 318 334 CATALYTIC (POTENTIAL).
FT DOMAIN 1013 1029 CATALYTIC (POTENTIAL).
FT DOMAIN 515 530 GLY-RICH.
FT DOMAIN 569 602 GLY-RICH.
FT DOMAIN 1278 1297 GLY-RICH.
FT DOMAIN 1767 1810 GLY/SER-RICH.
FT DOMAIN 2025 2040 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 2200 2241 GLN-RICH.
FT METAL 280 280 MAGNESIUM 1 AND 2 (BY SIMILARITY).
FT METAL 281 281 MAGNESIUM 2 (VIA CARBOXYL OXYGEN) (BY
SIMILARITY).
FT METAL 324 324 MAGNESIUM 1 AND 2 (BY SIMILARITY).
FT CARBOHYD 800 800 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 807 807 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MUTAGEN 1026 1026 G->R: ABOLISHES CATALYTIC ACTIVITY.
SQ SEQUENCE 2248 AA; 248899 MW; E459C718EB018668 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 2248;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANGQ 8
Db 1269 NGNGANG 1276

RESULT 11
CYPH_UROFA STANDARD; PRT; 163 AA.
ID CYPH_UROFA 000060;
AC 000060;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptidy-prolyl cis-trans isomerase (EC 5.2.1.8) (PPIase) (Rotamase)
DE (Cyclophilin) (Cyclosporin A-binding protein) (Planta-induced rust
protein 28).
GN PIG28.
OS Uromyces fabae (Rust fungus).
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Urediniomycetidae;
OC Uredinales; Pucciniae; Uromyces.
OX NCBI_TaxID=55588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12; TISSUE=Haustorium;
RX MEDLINE=97294919; PubMed=9150592;
RA Hahn M., Mendgen K.;
RT "Characterization of in planta-induced rust genes isolated from a
haustorium-specific cDNA library.";
RL Mol. Plant Microbe Interact. 10:427-437(1997).
CC -1- FUNCTION: PPIases accelerate the folding of proteins. It catalyzes
the cis-trans isomerization of proline imidic peptide bonds in
oligopeptides.
CC -1- CATALYTIC ACTIVITY: Peptidyproline (omega=180) = peptidyproline
(omega=0).
CC -1- ENZYME REGULATION: Binds cyclosporin A (CSA). CSA mediates some of
its effects via an inhibitory action on PPIase.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- DEVELOPMENTAL STAGE: Haustoria and rust-infected leaves. Also
observed, in lower levels, in spores or hyphae formed in vitro.
CC -1- SIMILARITY: Belongs to the cyclophilin-type PPIase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U81793; AAB39880.1; -.
DR HSSP; P05092; 2CPL.
DR InterPro; IPR002130; CSA_PPIase.
DR Pfam; PF00160; pro_isomerase; 1.

```

```

DR PRINTS; PR00153; CSAPPISMASE.
DR PROSITE; PS00170; CSA_PPIASE_1; 1.
DR PROSITE; PS50072; CSA_PPIASE_2; 1.
DR Cyclosporin; Isomerase; Rotamase; Multigene family.
SQ SEQUENCE 163 AA; 17992 MW; A00A641CBA0DBD1D CRC64;

Query Match 70.0%; Score 35; DB 1; Length 163;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGNGANGQXV 10
Db 67 NGNGANGBEST 76

RESULT 12
NIPL_MOUSE STANDARD; PRT; 218 AA.
ID NIPL_MOUSE 092287;
AC 092287;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Bcl2/adenovirus F1B 19-kDa protein-interacting protein 3 like (NIP3L)
DE (NIP3-like protein X).
GN NIP3L OR NIX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=99084982; PubMed=9867803;
RA Chen G., Cizeau J., Vande Velde C., Park J.H., Bozek G., Bolton J.,
RA Shi L., Dubik D., Greenberg A.;
RT "Nix and Nip3 form a subfamily of pro-apoptotic mitochondrial
proteins.";
RL J. Biol. Chem. 274:7-10(1999).
CC -1- FUNCTION: Induces apoptosis. Interacts with viral and cellular
anti-apoptosis proteins. Can overcome the suppressors BCL-2 and
BCL-XL, although high levels of BCL-XL expression will inhibit
apoptosis. May function as a tumor suppressor.
CC -1- SUBUNIT: Homodimer (Probable).
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- PTM: Undergoes progressive proteolysis to an 11 kDa C-terminal
fragment, which is blocked by the proteasome inhibitor
lactacystin.
CC -1- SIMILARITY: Belongs to the NIP3 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF067395; AAD03588.1; -.
DR MGD; MGI1332659; Bnip3L.
DR Apoptosis; Transmembrane; Mitochondrion.
FT TRANSMEM 187 207 POTENTIAL.
SQ SEQUENCE 218 AA; 23766 MW; EAA639DFDE50D7 CRC64;

Query Match 70.0%; Score 35; DB 1; Length 218;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANG 7
Db 48 NGNGANG 54

RESULT 13

```

NIP3L_HUMAN
 ID NIP3L_HUMAN STANDARD; PRT; 219 AA.
 AC O60238;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE BCL2/adonovirus E1B 19-kDa protein-interacting protein 3 like (NIP3L)
 DE NIP3-like protein X (BCL2/adonovirus E1B 19-kDa protein interacting protein 3A).
 GN NIP3L OR BNIP3A OR NIX OR BNIP3H.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCB1_TaxID=9606;
 RX MEDLINE=98183799; PubMed=9523198;
 RA Matsushima M., Fujiwara T., Takahashi E., Minaguchi T., Eguchi Y., Tsujimoto Y., Suzumori K., Nakamura Y.,
 RT "Isolation, mapping, and functional analysis of a novel human cDNA (BNIP3L) encoding a protein homologous to human NIP3.",
 RL Genes Chromosomes Cancer 21:230-235(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yasuda M., Han J.-W., Dionne C.A., Boyd J.M., Chinnadurai G.,
 RT "BNIP3A, a human homolog of pro-apoptotic protein BNIP3, promotes apoptosis and interacts with viral and cellular anti-apoptosis proteins".
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Fetal liver;
 RX MEDLINE=99084982; PubMed=9667803;
 RA Chen G., Cizeau J., Vande Velde C., Park J.H., Bozek G., Bolton J., Shi L., Dubik D., Greenberg A.,
 RT "Nix and Nip3 form a subfamily of pro-apoptotic mitochondrial proteins".
 RL J. Biol. Chem. 274:7-10(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Dermal papilla;
 RA Farooq M., Kim M.K., Kim Y.H., Seo J.M., Lee H.M., Sohn M.Y., Hwang S.Y., Chung H.-J., Im S.U., Jung E.J., Kim J.C.,
 RT "A catalog of genes in the human dermal papilla cells as identified by expressed sequence tags".
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Glactina M., Lee Y.-T., Miller J.L.,
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Bone marrow, and Cervix;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Stauber R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stetson M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uegin T.B., Toshiyuki S., Carinici P., Prange C., Rana S.S., Loguailano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W., Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -! FUNCTION: Induces apoptosis. Interacts with viral and cellular anti-apoptosis proteins. Can overcome the suppressers BCL-2 and BCL-XL, although high levels of BCL-XL expression will inhibit apoptosis. May function as a tumor suppressor.
 CC -! SUBUNIT: Homodimer (Probable).
 CC -! SUBCELLULAR LOCATION: Mitochondrial.
 CC -! PTM: Undergoes progressive proteolysis to an 11 kDa C-terminal fragment, which is blocked by the proteasome inhibitor lactacystin.
 CC
 CC -! SIMILARITY: Belongs to the NIP3 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AB004788; BAA28692.1; -
 DR EMBL: AF079221; AAC27723.1; -
 DR EMBL: AF067336; AAD03589.1; -
 DR EMBL: AF255051; AAF70290.1; -
 DR EMBL: AF452712; AAL50978.1; -
 DR EMBL: BC001559; AAH01559.1; -
 DR EMBL: BC009603; AAH09603.1; -
 DR GeneW: HGNC:1085; BNIP3L.
 DR MIM: 605368; -
 DR GO: GO:0005739; C:mitochondrion; TAS.
 DR GO: GO:0006917; P:induction of apoptosis; TAS.
 DR GO: GO:0006934; P:negative regulation of survival gene products; TAS.
 KW Apoptosis; Transmembrane; Mitochondrion.
 FT TRANSMEM 188 208 POTENTIAL.
 SQ SEQUENCE 219 AA; 23930 MW; 19372B897BC3609 CRC64;
 QY 1 NGNGANG 7
 DB 49 NGNGKNG 55
 Query Match 70.0%; Score 35; DB 1; Length 219;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 RESULT 14
 NTPA_PEA STANDARD; PRT; 455 AA.
 ID NTPA_PEA
 AC P52914;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nucleoside-triphosphatase (EC 3.6.1.15) (Nucleoside triphosphate phosphohydrolase) (NTPase) (Apyrase).
 OS Pisum sativum (garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
 NX NCB1_TaxID=3888;
 RX MEDLINE=96197404; PubMed=8616230;
 RA Hsieh H., Tong C.G., Thomas C., Roux S.J.,
 RT "Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin-associated NTPase in pea".
 RL Plant Mol. Biol. 30:135-147(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=cv. Alaska; TISSUE=Stem;
 RA Shibata K., Abe S., Davies E.,
 RT "Structure of the coding region and mRNA variants of the apyrase from Pisum sativum."

```

RL Acta Physiol. Plant. 20:3-13 (2001).
CC -!- FUNCTION: Might be involved in RNA transport out of nuclei.
CC -!- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z32743; CAAB3655.1; -.
DR EMBL; AB027614; BAB18900.1; -.
DR PIR; S65147; S48859.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39_1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
DR HydroLase; Nuclear protein.
SQ SEQUENCE 455 AA; 50PDP0023ABC4299 CRC64;

Query Match          70.0%; Score 35; DB 1; Length 455;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGNGANGQ 8
DB 322 NGGGGNGQ 329

RESULT 15
TBG RETFI          STANDARD; PRT; 478 AA.
AC P54405;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tubulin gamma chain (Gamma tubulin).
OS Reticulomyxa filosa.
OC Eukaryota; Granuloreticulosea; Athalamae; Reticulomyxidae;
OC Reticulomyxa.
OC NCBI_TaxID=46433;
OX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97273248; PubMed=9127728;
RA Kube-Grandgerath E., Schliwa M.;
RT "Unusual distribution of gamma-tubulin in the giant fresh water
RT amoeba Reticulomyxa filosa."
CC -!- FUNCTION: Tubulin is the major constituent of microtubules. Gamma
CC tubulin is found at microtubule organizing centers (MTOC) such as
CC the spindle poles or the centrosome, suggesting that it is
CC involved in the minus-end nucleation of microtubule assembly.
CC -!- SIMILARITY: Belongs to the tubulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X97250; CAAB5885.1; -.
DR InterPro; IPR008280; Tub_FtsZ_C.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_FtsZ.
DR Pfam; PF00091; tubulin; 1.
DR Pfam; PF03953; tubulin_C; 1.
DR PRINTS; PR01161; TUBULIN.
DR PROSITE; PS00227; TUBULIN; 1.

```

```

KW Microtubule; GTP-binding.
FT NP_BIND 141 147
SQ SEQUENCE 478 AA; 53558 MW; 4702908E7AEBB98C CRC64;

Query Match          70.0%; Score 35; DB 1; Length 478;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNGANGQ 8
DB 463 NGNGANGK 470

```

Search completed: July 21, 2004, 17:06:53
Job time : 2.34261 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2004, 17:02:32 ; Search time 1.52034 Seconds
(without alignments)
2075.310 Million cell updates/sec

Title: US-10-018-604-4
Perfect score: 50
Sequence: 1 NNGNNGQXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	90.0	630	10 Q40161	Q40161 lycopersio
2	43	86.0	720	16 Q8EXL7	Q8EXL7 leptospira
3	40	80.0	464	9 Q9AZ86	Q9AZ86 lactobacilli
4	39	78.0	137	16 Q7U161	Q7U161 mycobacteri
5	39	78.0	150	16 Q53843	Q53843 mycobacteri
6	39	78.0	487	2 Q50901	Q50901 myxococcus
7	38	76.0	427	5 Q94534	Q94534 drosophila
8	38	76.0	538	5 Q8NOR6	Q8NOR6 turbo marmo
9	38	76.0	913	12 Q39793	Q39793 equine aden
10	38	76.0	1051	16 Q8XYR0	Q8XYR0 ralsstonia s
11	38	76.0	1059	3 Q9C2U8	Q9C2U8 gibberella
12	38	76.0	2646	5 Q8I220	Q8I220 plasmodium
13	37	74.0	135	13 Q7M137	Q7M137 brachydanio
14	37	74.0	149	10 Q94CH0	Q94CH0 zea mays (m
15	37	74.0	350	16 Q82VL7	Q82VL7 nitrosomona
16	37	74.0	373	5 Q9U2I7	Q9U2I7 caenorhabdi

17	37	74.0	381	16 Q8E367	Q8E367 shewanella
18	37	74.0	568	5 Q9ML38	Q9ML38 pinctada ma
19	37	74.0	582	5 Q46042	Q46042 drosophila
20	37	74.0	582	5 Q8SY46	Q8SY46 drosophila
21	37	74.0	595	5 Q9W407	Q9W407 drosophila
22	37	74.0	676	5 Q9YDP7	Q9YDP7 drosophila
23	37	74.0	694	16 Q8YR19	Q8YR19 arabidena sp
24	37	74.0	807	5 Q9VNG2	Q9VNG2 drosophila
25	37	74.0	884	16 Q89L52	Q89L52 bradyrhizob
26	37	74.0	1015	16 Q8NU77	Q8NU77 staphylococ
27	37	74.0	1038	16 Q99RD2	Q99RD2 staphylococ
28	37	74.0	1247	5 Q962M1	Q962M1 plasmodium
29	37	74.0	1519	2 Q4E237	Q4E237 helicobacte
30	37	74.0	2327	5 Q9W0Y8	Q9W0Y8 drosophila
31	37	72.0	65	12 Q85391	Q85391 variola maj
32	37	72.0	65	12 Q89545	Q89545 variola vir
33	37	72.0	101	12 Q9QNH9	Q9QNH9 variola min
34	36	72.0	170	5 Q88IL7	Q88IL7 giardia lam
35	36	72.0	171	5 Q25756	Q25756 plasmodium
36	36	72.0	242	12 Q91FB0	Q91FB0 chilo tride
37	36	72.0	247	16 Q8G3M9	Q8G3M9 caenorhabdi
38	36	72.0	253	5 Q17426	Q17426 caenorhabdi
39	36	72.0	253	5 Q17536	Q17536 caenorhabdi
40	36	72.0	263	16 Q9L154	Q9L154 streptomyce
41	36	72.0	275	16 Q8FQ47	Q8FQ47 corynebacte
42	36	72.0	287	9 Q7Y4X8	Q7Y4X8 bacterioph
43	36	72.0	288	12 Q8UJ41	Q8UJ41 phage arl.
44	36	72.0	296	5 Q9V805	Q9V805 drosophila
45	36	72.0	403	16 Q8A0G5	Q8A0G5 bacteroides

ALIGNMENTS

RESULT 1

Q40161

AC Q40161

DT 01-NOV-1996 (T

DT 01-NOV-1996 (T

DE 01-JUN-2003 (T

DE POLYGALACTURON

OC Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC lamids; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4081;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93005745; PubMed=1392611;

RA Zheng L., Heupel R.C., Dellapenna D.;

RT "The beta subunit of tomato fruit polygalacturonase isoenzyme 1:

RT isolation, characterization, and identification of unique structural

RT features";

RL Plant Cell 4:1147-1156 (1992).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=VENT CHERRY;

RA Watson C.F., Schuchman B., Liu J., Dellapenna D.;

RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; W98466; AAA34181.1; -.

DR EMBL; U63374; AAB39547.1; -.

DR PIR; J01670; J01670.

DR InterPro; IPR004873; BURP.

DR Pfam; PF03181; BURP; 1.

KW signal.

FT SIGNAL

FT CHAIN

SQ SEQUENCE

Query Match

Best Local Similarity

Score 45; DB 10; Length 630;

Q8E367 shewanella
Q9ML38 pinctada ma
Q46042 drosophila
Q8SY46 drosophila
Q9W407 drosophila
Q9YDP7 drosophila
Q8YR19 arabidena sp
Q9VNG2 drosophila
Q89L52 bradyrhizob
Q8NU77 staphylococ
Q99RD2 staphylococ
Q962M1 plasmodium
Q4E237 helicobacte
Q9W0Y8 drosophila
Q85391 variola maj
Q89545 variola vir
Q9QNH9 variola min
Q88IL7 giardia lam
Q25756 plasmodium
Q91FB0 chilo tride
Q8G3M9 caenorhabdi
Q17426 caenorhabdi
Q17536 caenorhabdi
Q9L154 streptomyce
Q8FQ47 corynebacte
Q7Y4X8 bacterioph
Q8UJ41 phage arl.
Q9V805 drosophila
Q8A0G5 bacteroides

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
DB 243 NNGANGSEFV 252

RESULT 2

Q8EXL7 PRELIMINARY; PRT; 710 AA.
AC Q8EXL7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Putative TonB-dependent outer membrane receptor protein.
GN LBI91.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011607; AAM51750.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00531; TonB_boxC.
DR Pfam; PF00593; TonB_dep_Rec; 1.
KW Receptor; Complete proteome.
SQ SEQUENCE 710 AA; 80884 MW; 98CF29D77C6182E CRC64;

Query Match 86.0%; Score 43; DB 16; Length 710;
Best Local Similarity 80.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
DB 67 NNGANGQIV 76

RESULT 3

Q9AZ86 PRELIMINARY; PRT; 464 AA.
AC Q9AZ86;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE O-f585.
GN ORF585.
OS Lactobacillus johnsonii prophage Lj928.
OC Viruses.
OX NCBI_TaxID=139872;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20455575; PubMed=10998330;
RA Desiere F., Pridmore R.D., Brusson H.;
RT "Comparative genomics of the late gene cluster from Lactobacillus phages."
RL Virology 275:294-305(2000).
DR EMBL; AF195902; AAK27935.1; -.
SQ SEQUENCE 464 AA; 50789 MW; 5FF0826D6EC594F5 CRC64;

Query Match 80.0%; Score 40; DB 9; Length 464;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
DB 282 NNGANG 288

RESULT 4

Q7U161 PRELIMINARY; PRT; 137 AA.
AC Q7U161;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE PE-PGRS family protein.
GN PE_PGRS12 OR MB0855.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248336; CAD93717.1; -.
KW Complete proteome.
SQ SEQUENCE 137 AA; 13095 MW; 0473571C8EB5A64 CRC64;

Query Match 78.0%; Score 39; DB 16; Length 137;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNGANGQ 8
DB 119 NNGANGQ 125

RESULT 5

OS3843 PRELIMINARY; PRT; 150 AA.
AC OS3843;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE PE-family protein (PE-PGRS family protein).
GN RV0832 OR MT0854 OR MTV043.24.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Kiroh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Stulton J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwim M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT Laboratory strains."
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022004; CAA17638.1; ALT_INT.
DR EMBL: AE006974; AAK45095.1; -.
DR FIR; H70811; H70811.
DR TIGR; MT0854; -.
DR TubercuList; Rv0832; -.
DR InterPro; IPR000084; PE_region.
DR Pfam; PF00934; PE: 1.
DR ProDom; PD001223; PE_region; 1.
DR Complete proteome.
SQ SEQUENCE 150 AA; 14587 MW; B5FPCA611E447C4 CRC64;

Query Match 78.0%; Score 39; DB 16; Length 150;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GNGANGQ 8
DB 132 GNGANGQ 138

RESULT 6
AC 050901 PRELIMINARY; PRT; 487 AA.
ID 050901
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Beta-1,4-glycanase.
GN CELA.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DK101;
RA Culler L.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DK101;
RA Barry S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: X76640; CAA54086.1; -.
DR HSSP; P26222; ITML.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR003305; CBM_Cent.
DR InterPro; IPR008979; Gal_bind_like.
DR InterPro; IPR001524; Glyco_hydro_6.
DR Pfam; PF02018; CBM_4_9; 1.
DR Pfam; PF01341; Glyco_hydro_6; 1.
DR PRINTS; PR00733; GLHYDRLASE6.
DR ProDom; PD003733; Glyco_hydro_6; 1.
DR PROSITE; PS00655; GLYCOSYL_HYDROL_F6_1; 1.
SQ SEQUENCE 487 AA; 51034 MW; 59E6472401517B50 CRC64;

Query Match 78.0%; Score 39; DB 2; Length 487;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Beat protein precursor.
GN BEAT-1A OR CG4846.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97133212; PubMed=8978609;
RA Fambrough D., Goodman C.S.;
RT "The Drosophila beaten path gene encodes a novel secreted protein that
RT regulates defasciculation at motor axon choice points.";
RL Cell 87:1049-1058(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo W., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celniker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam K.,
RA Jatali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Iasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo W., Pittman G.S., Pan S., Pollard J., Furi V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slater E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Strikas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: U67057; AAB07545.1; -.
DR EMBL: AE003415; AAF44982.1; -.
DR

```
DR EMBL; AE003649; AAF53500.1; -.
DR FlyBase; FBgn0013433; beat-1a.
DR GO; GO:0007415; P: defasciculation of motor neuron; IMP.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 427 BEAT PROTEIN.
SQ SEQUENCE 427 AA; 47917 MW; 26646C37D88B119E CRC64;

Query Match 76.0%; Score 38; DB 5; Length 427;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGNANGQXV 10
Db 250 HGNANGNHV 259

RESULT 8
Q8N0R6 PRELIMINARY; PRT; 538 AA.
AC Q8N0R6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nacerein.
OS Turbo marmaratus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Verigastropoda; Trochoidea; Turbinidae; Turbo.
OX NCBI_TaxID=176104;
RN [1]
RP SEQUENCE FROM N.A.
RA Yano M., Miyashita T., Miyamoto H.;
RT "Shell matrix Nacerein, a novel family of carbonic anhydrase, is
RT conserved in bivalve and gastropod."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB073680; BAB91157.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0004089; F: carbonate dehydratase activity; IEA.
DR GO; GO:0008233; F: peptidase activity; IEA.
DR GO; GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO; GO:0008270; F: zinc ion binding; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR GO; GO:0006730; P: one-carbon compound metabolism; IEA.
DR InterPro; IPR001148; Euk Coahnd.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR005089; Peptidase_S26.
DR Pfam; PF00194; catp_anhydrase; 1.
DR ProDom; PD000865; Euk_Coahnd; 2.
DR PROSITE; PS00237; G_PROTEIN_RECPR_FL_1; 1.
DR PROSITE; PS00761; SBASE I 3; 1.
SQ SEQUENCE 538 AA; 57586 MW; 406351AB6175F170 CRC64;

Query Match 76.0%; Score 38; DB 5; Length 538;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGNANGQ 8
Db 413 NGNANGH 420

RESULT 9
Q39793 PRELIMINARY; PRT; 913 AA.
AC Q39793;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
```

```
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hexon protein.
GN HEXON.
OS Equine adenovirus type 1.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=46916;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97372953; PubMed=9229008;
RA Reibel G.H., Studdert M.J.;
RT "Identification, cloning and sequence analysis of the equine
RT adenovirus 1 hexon gene."
RL Arch. Virol. 142:1193-1212(1997).
DR EMBL; L79855; AAB88062.1; -.
DR HSSP; P03277; IDHX.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR000736; Adeno_hexon.
DR Pfam; PF01065; Adeno_hexon; 1.
DR ProDom; PD002815; Adeno_hexon; 1.
SQ SEQUENCE 913 AA; 102531 MW; 9799FCT746585010 CRC64;

Query Match 76.0%; Score 38; DB 12; Length 913;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGNANGQXV 10
Db 136 NGNANGARDV 145

RESULT 10
Q8XYR0 PRELIMINARY; PRT; 1051 AA.
AC Q8XYR0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Probable bacteriophage protein.
GN RSC1698 OR RS02885.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RC MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunne S., Demange N.,
RA Gaspin C., Lavie W., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigler P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646066; CAD15400.1; -.
DR InterPro; IPR003961; FN_III.
DR SMART; SM00060; FN3; 2.
KW Complete proteome.
SQ SEQUENCE 1051 AA; 112665 MW; C5E63D3FD742B9 CRC64;

Query Match 76.0%; Score 38; DB 16; Length 1051;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGNANGQXV 10
Db 1020 NGNANGRRV 1029

RESULT 11
```

```

09C2U8
ID 09C2U8 PRELIMINARY; PRT; 1059 AA.
AC 09C2U8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE NAD+ dependent glutamate dehydrogenase.
GN GDH1.
OS Gibberella fujikuroi (Bakaneae and foot rot disease fungus) (Fusarium
moniliforme).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=5127;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ms67;
RA Tudzynski B.;
RT "Cloning and characterization of a Gibberella fujikuroi (Fusarium
fujikuroi) NAD-dependent glutamate dehydrogenase gene."
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ310444; CAC27837.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR006095; GLFV_dehydrog_C.
DR InterPro; IPR006097; GLFV_dehydrog_N.
DR Pfam; PF00208; GLFV_dehydrog_1.
DR Pfam; PF02812; GLFV_dehydrog_N_1.
DR PROSITE; PS00074; GLFV_DEHYDROGENASE; 1.
KW NAD.
SQ SEQUENCE 1059 AA; 118944 MW; C79B4061C066DD7 CRC64;

Query Match 76.0%; Score 38; DB 3; Length 1059;
Best Local Similarity 60.0%; Pred. No. 4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NNGANGXQV 10
Db 52 NNGANGXHRT 61

RESULT 12
081220
ID 081220 PRELIMINARY; PRT; 2646 AA.
AC 081220;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Erythrocyte membrane protein 1 (PEM1).
GN VAR.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Bertram M., Churcher C., Harris B., Harris D.,
Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
Feltwell T., Goble A., Goodhead I., Gilliam R., Hamlin N., Hance Z.,
Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
Radandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).

```

```

DR EMBL; AL034557; CAD49094.1; -.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008640; Hep_Hag.
DR InterPro; IPR004258; PEM1P.
DR Pfam; PF05658; Hep_Hag; 1.
DR Pfam; PF03011; PEM1P; 2.
SQ SEQUENCE 2646 AA; 298286 MW; E71F001A1D6077B CRC64;

Query Match 76.0%; Score 38; DB 5; Length 2646;
Best Local Similarity 75.0%; Pred. No. 1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNGANGQ 8
Db 573 NNGANGGE 580

RESULT 13
07T137
ID 07T137 PRELIMINARY; PRT; 135 AA.
AC 07T137;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE SI:B21G18.3 (Novel protein similar to vertebrate gliacolin (C10))
DE (Fragment).
SI:B21G18.3.
GN Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL627248; CAE17620.1; -.
KW Ribonucleoprotein.
FT NON TER 1
SQ SEQUENCE 135 AA; 14653 MW; 8245CE31BDE7657 CRC64;

Query Match 74.0%; Score 37; DB 13; Length 135;
Best Local Similarity 85.7%; Pred. No. 68;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
Db 90 NNGANGSG 96

RESULT 14
094CH0
ID 094CH0 PRELIMINARY; PRT; 149 AA.
AC 094CH0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Seven transmembrane protein M105 (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Briggs S.P., Simmons C.R.;
RT "Manipulation of mlo genes to enhance disease resistance in plants.";
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Devoto A., Hartmann A., Piffanelli P., Elliott C., Simmons C.R.,
Taramino G., Goh C.-S., Cohen F.E., Schulze-Lefert P., Panstruga R.;

```

RT "Molecular phylogeny and domain-specific co-evolution of the plant-specific seven transmembrane Mlo family.";
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL:AY029316; AAK38341.1; -.
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR GO: GO:0008219; P: cell death; IEA.
 DR InterPro: IPR004326; Mlo.
 DR Pfam: PF03094; Mlo; 1.
 KW Transmembrane.
 FT NON_TER 1
 SQ SEQUENCE 149 AA; 16260 MW; D037057B000EB9A4 CRC64;

Query Match 74.0%; Score 37; DB 10; Length 149;
 Best Local Similarity 85.7%; Pred. No. 76;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNGANG 7
 |||:|
 Db 96 NNGANG 102

RESULT 15

Q82VL7 PRELIMINARY; PRT; 350 AA.
 AC Q82VL7;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN NE1058.
 OS Nitrosomonas europaea.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
 OC Nitrosomonadaceae; Nitrosomonas.
 OX NCBI_TaxID=915;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19718 / IPO 14298;
 RX MEDLINE=22586410; Pubmed=12700255;
 RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
 RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
 RA Arciero D.M., Hommes N.G., Whitaker M.M., Arp D.J.;
 RT "Complete genome sequence of the ammonia-oxidizing bacterium and
 RT obligate chemolithoautotroph Nitrosomonas europaea.";
 RL J. Bacteriol. 185:2759-2773(2003).
 DR EMBL: BX321859; CAD84969.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 350 AA; 35890 MW; F33BB4F45BE1CA36 CRC64;

Query Match 74.0%; Score 37; DB 16; Length 350;
 Best Local Similarity 60.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NNGANGQXV 10
 |||:|
 Db 186 NNGANGRNSI 195

Search completed: July 21, 2004, 17:08:21
 Job time : 10.5203 secs